Looking beyond what doesn’t work

– Results, policy implications and lessons learned from two outcome evaluations the Prime for Life alcohol prevention program.

Bo Sandberg

Rapport 2013:1

Kriminologiska institutionen

Stockholms universitet
Looking beyond what doesn’t work
– Results, policy implications and lessons learned from two
outcome evaluations of the Prime for Life alcohol prevention program

©Bo Sandberg
Kriminologiska institutionen
Stockholms universitet

Rapport 2013:1
Kriminologiska institutionens rapportserie
Redaktör: Felipe Estrada

ISSN 1400-853X
ISBN 978-91-979847-4-4
Foreword

During my time as an undergraduate student at Stockholm University in the late 1990s I took a course in evaluation at the Department of Education lead by Johan Dovelius and Ylva Sundberg. I later came to write my Master’s thesis based on an evaluation of an alcohol prevention project for the City of Stockholm. In 1999 I was contacted by my former teacher Johan who offered me to come to an interview to work as a consultant for Faugert & Co Utvärdering. This was a small company specialised in teaching, supporting and performing evaluations, mainly for governmental agencies in Sweden. For four years I worked as an consultant with evaluation within different policy fields such as education, research and innovation policy and environmental protection to mention a few. I got to work with skilled evaluators, people and organisations commissioning using evaluations to improve their activities and policies. In four years as an evaluation consultant I got a huge experience of teaching and performing evaluation studies withinin different policy fields. It also gave me the opportunity to write and present papers at evaluation conferences, a research article and a textbook on evaluation written together with Sven Faugert.

In 2003 I became a research assistant at the Centre for Social Research on Alcohol and Drugs (SoRAD) at Stockholm University, and in 2004 I got commissioned to do two research evaluations of the alcohol prevention program Prime for Life. The requirement for these two evaluations was higher standards for evidence than I had been used to as a consultant. The time line for generating knowledge was also longer. Previously the purpose of my evaluations had been to produce useful information for program improvement and decision making, while as a researcher I set out to produce knowledge about how the world works. The contrast between these two approaches to evaluation is something that I think is important to be aware of, because it might help to better understand the sometimes frustrating gap between policy making and knowledge generated by research. I think that one of the advantages of this thesis is that I am trying to bridge this gap by combining my own empirical studies and their results with a more theoretical and critical approach to what kind of knowledge is being produced and how it affects policy formation and improvement: Part of what I have learned through my experience as an evaluator is to be modest about what can be expected both from policy initiatives and from evaluations of social activities. Also, I have learned to think about and to become an active advocate for what can be
expected of and learned from evaluations. Both as a professional and as a researcher, I use different evaluation approaches, questions and educational tools to enhance the utilisation of evaluations. This is an important framework for this thesis; my desire is not only to focus on my empirical data and their results, but rather to take a critical look at them and at the contexts in which my own, and similar studies, are commissioned and to ask what use they (might) serve in relation to the improvement of social policy.

There are a lot of people that I wish to thank! First, I would like to thank my wife Kelly for supporting me in my work, and for believing in what I do! I also wish to thank Sven Faugert who recruited me as an evaluation consultant back in 1999. Sven has given me a solid experience and confidence in planning, performing and teaching evaluation which I could not have received from anyone else! I would like to thank Professor Janne Flyghed at the Department of Criminology who, as my supervisor and friend, has supported me both “inside and outside” the academia. Maud Quist, Head of the Evaluation Unit at the Swedish Research Council, has given me the support to be able to finish this thesis without compromising my regular duties as an analyst at the evaluation unit. I would like to thank Anders Bergmark, Professor at the Department of Social Work for his valuable comments on this thesis and Lars Brännström, Associate Professor at the Swedish Institute for Social Research (SOFI) at Stockholm University who has given advice on how to improve the statistical analyses in my studies.

Above and beyond these people, there are several others to thank: Friends and family; former colleagues at the Center for Social Research on Alcohol and Drugs and at the Department of Criminology at Stockholm University; the professors at the Criminology Department at the University of Southern Maine in the U.S, who gave me the opportunity to teach criminology at their department and who encouraged me to continue working on this thesis. Thank you all!

_The Dude abides_

Bo Sandberg

Stockholm, March 2013
Contents

1. Introduction ........................................................................................................................................... 1

   1.2. Objectives and overview .................................................................................................................. 1

   1.2. Criminological relevance ................................................................................................................ 2

   1.3 Swedish alcohol policy ..................................................................................................................... 4

       1.3.1 Alcohol policy and the evidence based practice movement .................................................. 5

2. Alcohol prevention, evaluation and knowledge transfer .............................................................. 8

   2.1. Alcohol prevention .......................................................................................................................... 8

   2.2 Evaluation ....................................................................................................................................... 9

       2.2.1 A brief history of the development of the evaluation field .................................................... 9

   2.3 Evidence-based practice and evaluation ........................................................................................ 11

       2.3.1 Outcome evaluations .................................................................................................................. 12

       2.3.2 Systematic reviews .................................................................................................................... 14

   2.4 The steps from evidence to evidence-based policy and evidence-based practice.... 16

       2.4.1 Shortage of evidence .................................................................................................................. 17

       2.4.2 The transfer of evidence to policy makers and professionals ................................................ 18

3. Introduction to the intervention and my studies ........................................................................... 21

   3.1 The Prime for Life program ............................................................................................................ 21

       3.1.1 Target group ............................................................................................................................... 22

   3.2 Evaluations and evidence relating to the Prime for Life program .............................................. 23

   3.3 Shared features of my evaluations of Prime for Life ................................................................... 25

       3.3.1 Program theory for the Prime for Life Program ...................................................................... 25

       3.3.2 Outcome variables ...................................................................................................................... 29

       3.3.3 Questionnaire design and testing ............................................................................................... 32

       3.3.4 Statistical analysis ....................................................................................................................... 32

       3.3.5 Ethical considerations ................................................................................................................ 34
4. Study 1: Evaluation of the Prime for Life Program at Örebro University ................. 35

4.1 Introduction .................................................................................................................. 35

4.1.1 Study objectives and research questions................................................................. 36

4.2 Method .......................................................................................................................... 36

4.2.1 Interviews with teachers and staff at the department.................................................. 37

4.2.2 Interviews with students........................................................................................... 37

4.2.3 Questionnaires.......................................................................................................... 38

4.3 Quality of outcome data ............................................................................................. 40

4.4 Main outcome results of the evaluation ...................................................................... 42

4.4.1 Outcomes for risk consumers ................................................................................... 43

4.4.2 Comments regarding the outcome results ................................................................. 45

4.5 Results regarding the implementation of the Prime for Life Program ...................... 46

4.5.1. Department representatives’ experience of Prime for Life...................................... 46

4.5.2 Short-term program evaluation results and student experiences of the program .... 50

4.5.3 Alcohol culture at the department ........................................................................... 52

4.5.4 Did the program produce any changes in alcohol culture?....................................... 54

4.5.5 The possibility of implementing the program in a university setting ....................... 55

5. Study 2: Evaluation of Prime for Life in the Swedish Armed Forces. ....................... 57

5.1 Introduction .................................................................................................................. 57

5.1.1 Objectives and research questions ........................................................................... 57

5.2 Method .......................................................................................................................... 57

5.2.1 Questionnaires.......................................................................................................... 58

5.3 Quality of outcome data ............................................................................................. 59

5.3.1 The experiment group ............................................................................................. 60

5.3.2 The comparison group ............................................................................................ 62

5.3.3 Comparing the experiment and comparison groups at T1 ..................................... 63
7.2 The narrow outcome focus ................................................................. 88
7.3 The top-down perspective ................................................................. 90
7.4 What could I have done differently? ......................................................... 92
8. Concluding remarks................................................................................. 95
References.................................................................................................. 100
Appendix 1: The AUDIT-test ................................................................... 111
Appendix 2: Questionnaire (T2) Swedish Armed Forces ............................... 112
Appendix 3: Outcomes of Prime for Life, Swedish Armed Forces ............... 124
1. Introduction

In this chapter, I will describe the purpose of this thesis and give an overview of its content. Further, I will describe the criminological relevance of the thesis and will briefly describe Swedish alcohol policy and the evidence-based movement, which is a framework for integrating practice, education and research with the objective of stimulating intervention studies and implementing documented methods in social work and criminal policy.

1.2. Objectives and overview
The empirical data for this study are taken from two evaluations of an alcohol prevention program. I will present the results from these studies and discuss their shortcomings and benefits and relate them to the evidence-based policy movement. I should emphasise that my ambition in the following is not to primarily focus on presenting these two studies of alcohol prevention. Rather, what I want to do is to use my studies to look critically at what value they, and similar studies, provide in relation to the understanding and enhancement of interventions and social policy.

Stated briefly, the purpose of this thesis is to critically discuss what possibilities there are to develop social interventions and the role of evaluations in this context. I will use my own studies to exemplify the shortcomings and strengths of different designs and methodological approaches to the study of the value and merit of interventions.

In Chapter 2, I will outline what the term “alcohol prevention” refers to, define evaluation and give an overview of the development of the field of evaluation research since the 1950s. In particular I will focus on the role of evaluations in the pursuit of so-called evidenced-based practice. I will problematize the prevailing evaluation perspectives that govern evidence-based practice and also the path from evidence to the development of social practices.

Thereafter (in Chapter 3), I will describe the Prime of Life program that I have evaluated and will also describe the “status” of this program from an evidence-based practice perspective. Further, I will present the shared features of the two evaluation studies I have conducted,
which include a program-theory perspective and the majority of the data collection and analysis procedures. I will also report on the ethical considerations that guided the studies and the reporting of results.

In the two following chapters (4-5), I will present the two evaluations and their main results. First (Chapter 4) I will present the outcome and implementation evaluation of the Prime for Life program at a department at Örebro University, then the outcome evaluation of the Prime for Life program in the Swedish Armed Forces (Chapter 5).

In Chapter 6, I discuss the evaluations in terms of the different kinds of knowledge that they produced and what their implications might be for alcohol policy.

In Chapter 7, I will discuss the strengths and shortcomings of my evaluation studies in relation to the desired design of outcome studies in evidence-based practice and in a number of other respects. I will also summarise what I would do differently if I were to conduct my studies again.

In the final chapter (8), I will revisit the alcohol policy and evidence-based practice context that was described in the first chapters. I will finish by presenting what I think constitute necessary elements in evaluations of alcohol prevention initiatives in order for them to better support the development of effective social interventions.

1.2 Criminological relevance

The two studies that I will present in this thesis are evaluations of a program that targets alcohol consumption and risks associated with alcohol consumption. The program studied constitutes an example of social prevention aimed at preventing and reducing harms caused by alcohol consumption in society (see Chapter 2.1).

In Sweden there is a strong link between the total level of alcohol consumption within the population and levels of alcohol-related harm...” (Norström & Ramstedt 2006) and a positive relationship between alcohol and violence (Tryggvesson 2005). The criminologist Leif Lenke found that changes in per capita consumption were related to criminal violence in Sweden...
(Lenke 1989) and that the relationship was stronger in countries with drinking patterns that involved higher levels of drinking to the point of intoxication (such as Finland). Von Hofer (2011) has shown in longitudinal studies of Swedish crime statistics that, in general, there is an historical co-variation between alcohol consumption and changes in levels of assault, manslaughter and murder. By decreasing the total level of alcohol consumption in Sweden, the total number of violent crimes can be reduced (Lenke 1989); less severe but more frequent acts of violence can be most effectively reduced by actions targeting the public (i.e. restaurant) consumption of alcohol, whereas more severe violence (aggravated assault) is most effectively reduced by selective interventions aimed at alcohol- and drug abusers. On the basis of aggregated time series of alcohol consumption and police reports (1956–1994), Norström (1998) has shown that “…the assault rate is related to consumption of beer and spirits in bars and restaurants, while the homicide rate is linked to consumption of spirits in private contexts” (Norström 1998, 689). Haggård-Grann et al. (2006) showed in a clinical study of 133 violent offenders that alcohol was a strong trigger of criminal violence. Kühlhorn et al. (1984) and Romelsjö (1995) have shown that the majority of both perpetrators and victims of violence had been drinking at the time of the crime or victimization.

Andréasson et al. (2006) estimated that a 40 per cent tax cut on spirits and 15 per cent on wine in Sweden would lead to an estimated increase in total annual per capita alcohol consumption of 0.35 litres, which would cause 289 additional deaths, 1627 additional assaults and 1.6 million additional days of sick leave per year. In an ambitious effort to estimate the social costs of alcohol in Sweden in 2002 using the cost-of-illness methodology, Johansson et al. (2006) estimated the costs resulting from the consequences of alcohol-related crimes to amount to approximately 3.5 billion SEK annually.

Overall, there is a strong association between alcohol and social problems in Sweden. There are also numerous international studies indicating that alcohol abuse is closely associated with violence and other types of criminality and social problems (see for example Murdoch, Phil & Ross 1990, Holder 1997, Martin 2001, WHO 2007).\(^1\) WHO’s Expert Committee on Problems Related to Alcohol Consumption (WHO 2007) concluded that:

---

\(^1\) A summary table of alcohol and violent crime studies can be found in Murdoch, Phil & Ross 1990, Table 1.
(...) harmful use of alcohol has immediate and long-term effects impinging on every aspect of life, with alcohol intoxication impairing personal safety with its association with violence and other injuries, alcohol dependence as a substantial risk of regular heavy drinking, and toxic effects of alcohol potentially affecting all organs and systems of the body (WHO 2007, 23).

Given the strong association between alcohol consumption and social problems, it is important to study the value and merit of actions aimed at reducing the problems associated with alcohol consumption. As will be illustrated in the next chapter, there has been an increased demand for evaluations that measure the effectiveness of actions taken to reduce alcohol consumption and/or the harms associated with alcohol consumption.

1.3 Swedish alcohol policy

The main objective of Swedish alcohol policy is to reduce the total level of alcohol consumption in society (prop. 2005/06:30). Swedish alcohol policy is based on the total consumption model, which suggests that there is a positive relationship between the total level of alcohol consumption in a society and its level of alcohol-related problems; the model stipulates that alcohol-related harms and problems in society will be reduced if the total level of alcohol consumption is reduced.

Traditionally, Swedish alcohol policy has focused on high alcohol taxes, a state monopoly on the sale of alcohol and limited import quotas as ways to keep the total level of alcohol consumption down. With Sweden’s European Union membership (1995) all state monopolies, with the exception of retail sales, were abolished and since then alcohol taxes have been adjusted downwards while alcohol import quotas have been adjusted upwards (Norström & Ramstedt 2006). Total alcohol consumption in Sweden increased by about 30 percent from 8.0 litres of pure ethyl alcohol per capita in 1996 to 10.5 litres in 2004 when my studies were initiated (Andréasson, Nilsson & Bränström 2010, Ramstedt et al. 2010). In 2010, total consumption was about 9 litres per capita (Ramstedt et al. 2010).
The foundation of Swedish national alcohol policy has changed dramatically since the mid-1990s and policies have been re-directed from using political instruments to keep total consumption down towards increased individual responsibility and a focus on local prevention measures. The national alcohol plan of 2006 (prop.2005/06:30) had a strong focus on local alcohol prevention and on vigorously promoting the development of prevention at the municipal level (Sandberg 2011). The increased focus on local alcohol prevention in Sweden is also closely associated with the government’s push towards a focus on local crime prevention in the mid-1990s, when the Swedish government stressed the importance of crime prevention at the local level by means of collaborations between organizations, private companies and local administrations. For the fight against criminality to have an impact at the local level ‘(…) it is necessary that all good forces collaborate and that the struggle is conducted on the basis of broad alliances’ (Ds 1996:59, 55. Author’s translation). The government encouraged municipalities to establish local crime prevention councils. In 1999 the Swedish National Council for Crime Prevention (Brå) was given the responsibility of supporting and promoting local crime prevention councils. By 2002, 202 out of 289 Swedish municipalities had initiated some sort of collaboration or council for local crime prevention (Brå 2002) and by 2005 there were around 290 local crime prevention councils in Sweden (Brå 2005). As it turns out, the most common focus for local crime prevention efforts has been alcohol- and drug prevention among youths (Brå 2005).

1.3.1 Alcohol policy and the evidence based practice movement

Goal orientation and intensified efforts to assess the results and outcomes of public policy have been intensified in Swedish public policy over the past 25 years or so (Tarschys & Lemne 2012). Since the 1980s, public management in Sweden has been geared towards goal-oriented thinking, competition, outsourcing and privatization (Vedung 2010). Consequently, there has been an increased focus on control and auditing as a means of ensuring that public services are efficient (Reeder 2010). This reform of government policies in order to modernise and render the public sector more efficient is in line with the tenets of the New Public Management (Ellefsen 2011).
Evidence-based practice (EBP) constitutes an effort to create a framework for the integration of practice, education and research with the objective of stimulating intervention studies and implementing scientifically documented methods in e.g. social and criminal policy (see Anderberg & Dahlberg 2005). Alongside the shift in national alcohol policy from control measures towards increased individual responsibility and local prevention, there has also been an increased focus on scientific evidence in social policy in Sweden (Bergmark 2005).

The methodological inspiration of EBP comes from Evidence Based Medicine (EBM), which aims to apply the best available evidence to clinical decision making. At the end of the 20th century, evidence-based practices spread from the field of healthcare to new areas such as social care, criminal justice and education (Walshie & Rundall 2001). The concept of EBP in social policy in Sweden can be dated to the early 1990s, when the Centre for Evaluation of Social Services (CUS) was established at the National Board of Health and Welfare (in 1992). However, it was not until the late 1990s and early 21st century that the concept of evidence-based social policy really became established (Bergmark & Lundström 2006).

The concept of EBP, whereby policy decisions, including decisions relating to alcohol prevention and treatment (von Greiff 2008), should be guided by the best scientific evidence, has become a key part of social policy, including alcohol, drug and crime prevention, in Sweden. The goal for national social policy, for example, is to develop evidence-based practice (SOU 2008:18), which includes strategies to tackle individuals with alcohol or drug problems. In 2007, the Swedish government established the Social Council to support the government in establishing evidence-based social policy in Sweden 2008-2010 (Dir. 2007:161). Further, the Swedish National Council for Public Health recommends evidence-based methods for preventive work to combat alcohol, narcotics and tobacco related problems (FHI 2009). The use of “effective prevention methods” has also been endorsed through the encouragement of municipalities to develop alcohol prevention initiatives (FHI 2008). The Swedish Prison and Probation Service has evaluations of effective programs (what works) as a prioritized area for research and development work in relation to treatment programs (Kriminalvården 2007).

In this thesis the term evidence-based practice encompasses both policy-making and direct practice (see Mullen & Streiner 2004).
The rationale for the focus on evidence-based policy is the idea that there is inefficiency and a lack of research-based policies. Within the evidence-based practice movement\(^3\) there is a desire to bridge the gap between research and practice in the field of social policy (see for example Sundell et al. 2009), and ideas about what constitutes “scientific evidence” have been adopted from evidence-based medicine. Strategies and interventions that have been tested and “proven” to be efficient in outcome evaluations, preferably using randomized controlled experimental designs, should be guiding policy-making and practice (SOU 2008:18, see also Socialstyrelsen 2009b, 2010b). I will return to the role of evaluations in EBP in Chapter 2, below.

\(^3\) The term “evidence-based practice movement” refers to the voices that insist that policy and practice must be based on research evidence (Hammersley 2005, Marks 2002)
2. Alcohol prevention, evaluation and knowledge transfer

In this chapter I will define alcohol prevention and evaluation. I will describe how the field of evaluation research has developed and the role of evaluation in EBP, and I will also discuss the most frequent evaluation approaches used as an input to EBP in more detail. I will also illustrate and problematize a number of issues regarding evidence production and the transfer of evaluation results to policy makers and professionals.

2.1. Alcohol prevention

Prevention (from the Latin word *praevinio*) is about hindering something from occurring. Alcohol prevention usually refers to strategies to reduce alcohol-related harms either by means of situational prevention, e.g. by increasing taxes on alcohol, or social prevention, e.g. by changing people’s attitudes towards alcohol (Socialstyrelsen 2009a).

In alcohol prevention, three different levels of prevention are usually referred to (see for example FHI 2004):

1. Primary prevention – refers to activities that are aimed at the general population and undertaken before a problem occurs in order to prevent problems from arising in the first place. Most educational alcohol prevention programs fit into this category.
2. Secondary prevention – refers to activities that are aimed at individuals or groups that are at risk for alcohol-related problems.
3. Tertiary prevention – refers to activities that are aimed at individuals who already have identified problems. This might for example involve interventions such as treatment programs for alcohol abusers.

Sometimes the terms universal (i.e. primary) prevention, selective (i.e. secondary) prevention and indicated (i.e. tertiary) preventions are used (Eriksson Tinghög 2013).
2.2 Evaluation

Evaluation is a common concept in education, research and development in most areas of the public sector. What it means to evaluate varies across different settings. The classic definition of evaluation provided by the evaluation theorist Michel Scriven describes evaluation as “the process of determining the merit, worth and value of things” (Scriven 1991, 1). I define evaluation as a systematic inquiry into an activity’s value and meaning, which is similar to the definition of evaluation employed by many Swedish governmental agencies and organizations (see for example Integrationsverket 2000, Statskontoret 2001, Svenska Missionsrådet 2003, Vetenskapsrådet 2011). Key to this definition is the view that evaluation involves a systematic and structured process that applies scientific methods to gather data about activities. An evaluation studies a specific activity, such as a project, an intervention, a program or a policy, and approaches its meaning and value by assessing it on the basis of one or more specified criteria.

2.2.1 A brief history of the development of the evaluation field

The roots of evaluation go back hundreds of years. However, its establishment as a scientific tradition and professional practice is a relatively recent phenomenon (Rossi, Lipsey & Freeman 2004). Massive federal programs aimed at curbing poverty, starvation and unemployment were implemented in the U.S. after World War II (the Great Society social reforms), and as public resources were poured into these programs, there was increasing pressure to provide evidence of their effectiveness (Patton 1997).

In the 1950s and 1960s, social scientists were engaged to perform evaluations of social activities such as delinquency prevention programs, psychotherapeutic and psychopharmacological treatment programs, educational programs etc. The focus was directed at quantitative and experimental studies. Evaluations served the purpose of guiding decision making by analysing which kind of activities tax money should be used for. What
was looked for in evaluations was “the knowledge of results” (Rossi, Lipsey & Freeman 2004, 8).

In the 1970s, evaluation began to take root in different European countries. Evaluation became established as a research area within the social sciences and the evaluation field became more diverse in its methods of data collection and analysis, which led to the emergence of new approaches to the study of the value and meaning of social activities. During the 1970s, a number of scientific journals focused on evaluation were initiated (Hogan 2007, Rossi, Lipsey & Freeman 2004), and universities began to offer courses in evaluation methodology (Hogan 2007). By the late 1970s the interest in evaluation had become so great that professional organizations for evaluators had also been established (Hogan 2007, Patton 1997).

During the 1970s and 1980s, the issue of how results were produced became a focus of special interest (Karlsson 1999) along with that of how evaluation results were used. By the early 1980s, the evaluation field had developed beyond being centred on the use of experimental methods into a methodologically diverse field of research. Qualitative methods gained ground as researchers sought to develop a better understanding of the activities that were being evaluated (Morris, Fitz-Gibbon & Lindheim 1988). A new role emerged for evaluators – to contribute to improving the activities that were being evaluated (Patton 1997). Evaluation also grew from being a field dominated by researchers, to become a political activity and an integrated part of the development of social policy and public management (Sandberg & Faugert 2012). During the 1980s and 1990s, evaluations that focused on participation and dialogue became more common. Fourth Generation Evaluation (Guba & Lincoln 1989) is one important example of the way in which the evaluation field developed during this period. In this type of evaluation, the concerns and questions of different stakeholders⁴ constitute the focal point and guide the evaluation.

In the early 1990s, there was also a trend towards theory driven evaluations. This type of evaluation focuses on the theoretical foundations of social actions, and attempts to improve knowledge about how effective activities work. This perspective represents a contrast to

---
⁴ The term stakeholder is both here and in the rest of the text to be understood its widest sense, i.e. funders, citizens, decision-makers, governmental agencies and local authorities, participants in an activity, the target group of an activity and others affected by it (Guba & Lincoln 1989, Vedung 2002).
traditional method-oriented outcome evaluations, which according to the American sociologist Huey-Tsyh Chen, involve conducting experimental studies in a mechanical fashion (Chen 1994).

At the end of the 20th century, there was an increased focus on fiscal conservatism in both the U.S. and Europe, which lead to the requirement of studying the financial and social impact of social programs (Rossi, Lipsey & Freeman 2004).

Evaluation is today a research field that spans over different scientific disciplines and involves a variety of perspectives and methodological approaches (Sandberg & Faugert 2012). The development of evaluation in Sweden has been rather similar to that witnessed in the U.S., although in Sweden developments have been more concentrated to the last decades of the 20th century (Gröjer 2004) – up until the 1970s, the term “utvärdering” (i.e. evaluation) was seldom or never used. During the 1980s, evaluation started to become more common both as a concept and an activity (Karlsson 1999), and the level of interest in evaluation increased significantly during the 1990s (Lindgren 2008).

2.3 Evidence-based practice and evaluation

Since the mid-1990s, the evidence-based practice movement has had a significant impact in certain policy areas in Sweden. Within the field of social policy, EBP tends to be associated with guidelines, systematic reviews of the effects of interventions, the implementation of validated tools for assessing clients and the development of evaluation systems (von Greiff 2008). There is a strong desire to create a more distinct relationship between needs, interventions and the results of social interventions.

Essential to EBP is a need to assess and rank available research so that only results from evaluation studies that meet the requirements for EBP are used to guide practice (Reynolds 2000). Although there are different ideas within EBP about how different kinds of research should be assessed, there is a general preference for randomized control studies, which are sometimes referred to as the “golden standard” of evaluation (using a terminology adopted from the bio-medical field of research).
“The evidence-based tradition recommends the most rigorous evaluation methods, favours quantitative methods and warns against using unstructured qualitative methods” (Ellefsen 2011, 112). The basic idea is that by using a certain set of scientific methods, social researchers can calculate what works and what does not work in the field of alcohol and crime prevention, and through knowledge of what works, policies and practices will improve. To assess “what works”, outcome evaluations based on experimental designs are desirable, since these give the researcher a certain control over the conditions when studying whether or not a certain intervention is better than either no intervention or a different intervention (SOU 2008:18).

Outcome evaluations and systematic reviews are two evaluation perspectives that are essential to EBP, and I will therefore briefly describe these below.

2.3.1 Outcome evaluations

Evaluations that measure the outcome of interventions are essential to the assessment of what works and what does not work in alcohol prevention. Outcome evaluations preferably use an intervention’s own goal(s) as the focus for the evaluation, since in a democracy it is of major interest to find out whether politically-decided actions produce the outcomes that have been promised (Vedung 2009). Since political and administrative goals are almost always vaguely formulated (Vedung 2009) and are quite often inappropriate for use as the guiding principles of an evaluation (Patton 1997), the researcher often defines the desired results of an intervention. The evaluator sets up specific desired (or undesired) consequences of the intervention that he or she wants to study, regardless of whether these are included among the goals set by those responsible for the intervention (Sandberg & Faugert 2012).

From a statistical standpoint, the ideal design for an outcome evaluation is an experimental study that can provide statistically robust results that can be generalised. Experimental outcome evaluations often involve rigorous statistical quality criteria (American Evaluation
Association 2004, Biemer & Lydberg 2003, SBU 2001). Ideally, a study is set up as comparative study with pre- and post-tests, randomization, data with high specificity (i.e. standardized tests, medical observations or register studies), is replicable and uses outcome measures that target what the intervention aims to change (Socialstyrelsen 2010b). Proper randomization ensures that there are no systematic differences between intervention groups with regard to factors that may affect the outcome.

For various reasons (some of which will be further explored throughout this thesis), evaluations of social interventions are often designed as quasi-experiments. The results of an activity for an experiment group are then compared to a comparison group that has not been part of the activity (Dahmström 1996). The comparison group includes individuals (or units) that are as similar as possible to the experiment group on key variables. Since we can never be sure that all relevant variables have been controlled for between the groups, we cannot be sure whether non-random factors, such as motivation to participate in the intervention, are “disturbing” the possibility of drawing statistically robust conclusions (Dahmström 1996). On the basis of a quasi-experiment, we cannot know with the same statistical certainty as in the randomized experiment whether outcomes are due to the intervention or to differences between the experiment and comparison group. This limits the possibilities of making generalizations based on the results of the evaluation (Dahmström 1996, Sandberg & Faugert 2012).

Some remarks regarding randomized outcome studies

When discussing EBP and evaluations of social interventions, we also need to consider some of the limitations of randomized experiments:

1. A randomized experiment may restrict other types of activities for both the experiment and the control group. A randomized experimental design limits the choices for those implementing an intervention, since the evaluator wants the maximum control possible over the experiment (Sandberg & Faugert 2012).
2. It may be difficult, and unethical, to randomly assign individuals, groups or activities to experiment and control groups; it restricts the opportunities for participants and stakeholders to choose what feels relevant for them to participate in (Lilja et al. 2004). It may be unethical not to offer all clients a treatment if there are valid reasons for believing that this treatment is better than no intervention at all (Dahmström 1996).

3. In many circumstances, random allocation does not work at all; this might be the case for example when the unit of study is a community rather than an individual (Tilley 2000). In community interventions, randomization is also likely to undermine the social and collective mechanisms that are central to planning and implementing a “community effort” (Hope 2005).

2.3.2 Systematic reviews

In systematic reviews, the results from several experimental or quasi-experimental outcome evaluations are compiled and analysed (Lipsey 2000, Lipsey & Wilson 2000). A review usually begins by studying published scientific articles or research reports regarding a given type of intervention. Using specified methodological criteria, relevant evaluation studies are identified and collected. Randomized controlled studies are usually ranked highest (Pawson 2002).

Meta-analysis is a statistical approach used to combine the data derived from a systematic-review. In a meta-analysis, effect sizes (ES) are used to compare different studies and to calculate the mean effect of the type of intervention being studied. ES expresses the difference between two groups, independent of sample size and without using statistical significance. This means that different programs can be compared using a common scale regardless of the size of the studies (Coe 2000, Hojat & Xu 2004). The word “effect” is commonly used within EBP to describe an intervention’s impact, i.e. the outcome above and

---

5 The term “standardised mean difference” (SMD) is often used as a synonym for “effect size” (Deeks, et al. 2008)
beyond what would have occurred anyway without the specific activity (Sandberg & Faugert 2012).

A systematic review serves the purpose of gathering information from multiple studies about an intervention. Meta-analysis has received renewed attention as a result of the popularity of EBP, since systematic reviewers claim that meta-analysis is a more valid way of reviewing research than the use of a narrative review for example (Egger et al. 2000). It is a powerful tool for guiding policy makers in their choice of actions, since a meta-analysis can summarize positive outcomes based upon hundreds of evaluations and can identify characteristics that are central to how an intervention works (Lipsey 2000). In addition, from a statistical standpoint, the results presented provide reliable answers about how effective an intervention is.

Some remarks regarding systematic reviews

There are a number of disadvantages associated with systematic reviews that it is important to be aware of:

1. Publication bias, i.e. when published research is systematically unrepresentative of the completed studies, can severely affect the results of systematic reviews (Dickersin 2005, Rothstein 2007). The reporting of outcomes within individual studies is often incomplete and biased as a result of the selective non-publication of statistically non-significant results in scientific papers (Chan & Altman 2005).

2. One inaccurate or poorly conducted evaluation included in a meta-analysis may involve a risk that the entire analysis will be flawed. When calculating the mean effect size of studies in a meta-analysis, there is a risk for skewed results caused by randomized experiments of inadequate quality (Petrosino 2003, Sutton et al. 2001). On the other hand, inclusion quality criteria that are set too high may leave the meta-analysis with too small a sample.
3. In a systematic review, only a fraction of the gross number of all the published evaluations and studies on a specific type of intervention will be included. The meta-analysis may represent a large proportion of the available experimental outcome evaluations, but need not be representative of all the research-based evaluations conducted into the type of intervention being studied. To take one example, in a study of the treatment of alcohol- and drug-problems conducted by the Swedish Council on Health Technology Assessment (SBU), 478 articles regarding the treatment of risky alcohol consumption were identified. Of these, 27 (6%) were randomized controlled studies that were then included in the study (SBU 2001). This is obviously both a weakness and a strength; in a meta-analysis a lot of research findings are ignored because they do not meet the inclusion criteria, while at the same time the strict protocol creates a solid foundation for generalizations based upon methodologically similar studies.

2.4 The steps from evidence to evidence-based policy and evidence-based practice

The concept of EBP relies on the idea that systematic knowledge about evidence is being spread and used by decision-makers and professionals within the field of alcohol prevention and treatment. For EBP to work effectively, there are many requirements that need to be met and I will here briefly discuss two central conditions necessary for the transfer of a certain kind of scientific knowledge to the development of policies and practices. I will return to these issues in the final chapters (Chapters 7-8) where I will discuss some of the problems and consequences associated with the use of a narrow outcome perspective when studying alcohol prevention initiatives.

The National Board of Health and Welfare is the Swedish governmental agency that has been commissioned by the government to describe trends in public health and social conditions. The agency cooperates with SFI Campbell to produce reviews of the effects of interventions in the fields of social work and crime prevention. SFI Campbell is part of the international Campbell Collaboration that works to improve decision-making by means of
systematic reviews on the effects of interventions in the areas of education, crime and justice, and social welfare (Sandberg & Faugert 2012). Since the National Board of Health and Welfare is one of the main advocates of evidence-based social policy in Sweden, I will be using a number of the agency’s own studies to illustrate some of the problems associated with the transfer of evidence to policy-making and practice.

2.4.1 Shortage of evidence

Studies of the desired scientific quality need to be both available and used as a guide for evidence-based practice, since this requires evidence-based interventions. To determine whether an intervention should be labelled “evidence-based”, it needs to have been studied and proven to be effective in accordance with the scientific criteria discussed above.

In a study of PhD-theses published in the fields of public health science, nursing science, criminology, psychology, education science, social work and sociology during the period 1997–2006 (N=1409), the National Board of Health and Welfare concluded that 181 (13 %) of these PhD-theses studied the effects of interventions. Of the 181 studies, 56 had a design that met the desired quality criteria to serve as an input to EBP (Socialstyrelsen 2010b). The study concluded that “the low number of PhD-students who have worked with outcome evaluations is deemed to complicate the implementation of an evidence-based practice in social work” (Socialstyrelsen 2010b, 7. Author’s translation).

From the perspective of the National Board of Health and Welfare, it is problematic for the development of evidence-based practice that there is insufficient interest (or perhaps competence) to support the desired development of EBP. In 2010, the National Board of Health and Welfare reported that evidence-based methods did not have a major role in social policy and social work programs at Swedish universities: “This means that EBP is not structuring the student’s education but is instead at risk of becoming a disconnected element that is often only introduced towards the end of the program” (Socialstyrelsen 2010a, 15. Author’s translation).

The National Board of Health and Welfare maintains a list of Swedish outcome studies of social interventions which it updates on a quarterly basis. The list includes studies
(conducted after 1990) that compare at least two alternatives (i.e. that include an experiment and a control group) and that have studied the intervention over time (i.e. with pre- and post-tests). In the early fall of 2012, the list included only 21 studies focused on interventions that primarily dealt with alcohol consumption or alcohol-related problems.⁶ (Two of these studies are the evaluations that will be presented in Chapters 4-5.) In Sweden, the production of scientific studies focused on the outcomes of interventions is rather modest (Alexandersson 2006).

Since there are not enough outcome evaluations available in Sweden, EBP relies heavily on international studies. Today, thousands of international meta-analyses have been published (Shadish et al. 2005). It should however be noted that the generalizations made in relation to the population of a given experiment cannot automatically be transferred to another population. Systematic reviews from other countries might therefore have limited scientific value as guidance for interventions in Sweden (Lilja et al. 2004).

2.4.2 The transfer of evidence to policy makers and professionals

A big challenge, and perhaps the greatest one, is to transfer evidence to policy makers and professionals and to make sure that knowledge about “what works” is being put into practice. In order for interventions that have been proven to be effective to be put to use, decision-makers and professionals must receive research results in a format that gives them the opportunity to use them, and they also need to have the capability to effectively implement those interventions that have been proven to be effective.

When evidenced-based social policy was initiated by the government in Sweden, it stressed the importance of improving the opportunities for professionals to stay abreast of the research and to incorporate it into their professional role (see for example SOU 2008:18). The background to this is related to the fact that social work has a relatively short history as an academic profession in Sweden. In contrast to the medical profession, for example, social work was up until the early 1990s conducted by laymen (Alexandersson 2006). Consequently, the EBP movement, at least with regard to social work, has lead both to new

requirements for higher education in social work and a need for PhD-training for researchers in social work. The National Board of Health and Welfare has developed different tools to support social workers in finding information and knowledge based upon evidence (Alexanderssson 2006, Socialstyrelsen 2011).

The idea of evidence-based policy and practice as a driving force towards the more efficient management of investments, using a model of research that sets out to deliver conclusive answers about what works (Simons 2004), is an appealing one, since it creates a linear link between research evidence and policy making. Within evaluation research, however, the ideas of EBP have sometimes been met with scepticism as the development of the evaluation field has become more responsive to different values and interests in a pluralistic society. It is argued that when it comes to decision making, researchers are only one of many parties that may want to have a say (Weiss 1999). In the early 1960s, social and behavioural science thought that science could be put to work to provide solutions to social problems, and there was a dream that scientific perspectives would be taken into account in policy making. However, the complexity of decision-making systems limited the influence of evaluations (Sanderson 2000, Weiss 1999). “While all types of applied science suffered from underuse, non-use seemed to be particularly characteristic of evaluation studies” (Patton 1997, 8). Today, it is well recognized in evaluation research that evaluation results are seldom used in decision making, at least not in an instrumental way (see for example Feinsten 2002, Patton 1997, Sandberg & Faugert 2012, Statskontoret 2002b). Consequently, many evaluators have adopted a more modest and educative approach to the role of informing policy and practice (Simons 2004, see also Patton 1997).

A survey of directors of social services in 290 Swedish municipalities indicated a great deal of interest in evidence-based practice among social service directors (Socialstyrelsen 2011). At the same time, the study showed that the directors had limited knowledge about what the term evidence-based interventions actually meant. The study concluded that even if the conditions required to implement evidence-based practice within the social services in Sweden may be present, there is limited knowledge about the effectiveness of different methods – and further developments may be hampered by a lack of capability for the training and quality controls required to support the implementation of evidence-based methods and support systems (Socialstyrelsen 2011).
Some have argued that the realization of EBP is likely to be complicated by the fact that it ignores the kind of knowledge that is required by professionals in decision making. A myriad of factors are taken into account when professionals make decisions, e.g. in the form of the integrated experience and knowledge that enables professionals to use tacit knowledge to cope in complex and unstable situations (Simons 2004, see also Sanderson 2000).

Stated briefly, the key to the success of EBP is in large part dependant on how decision-makers and professionals view knowledge and research, as well as on the opportunities that are open to them (e.g. financial, political, competence- or interest-based) when choosing and implementing certain interventions. There is a lot of evidence to suggest that decision making in professional contexts is far from being the linear and rational relationship that is often assumed by the scientists, politicians and professionals who advocate the EBP movement. It should be noted however that the EBP movement is still quite young and that, over time, it might lead to the emergence of a policy context in which the behavioural and social sciences have more influence in providing solutions to social problems that are then adopted in a more instrumental sense by policy makers and professionals. I develop my own views on the possibilities of such a development in the final chapter (8) of this thesis.
3. Introduction to the intervention and my studies

In Chapters 4-5 I will present the results from my two evaluations of the Prime for Life program. In this chapter I will describe the intervention, as well as the shared features of my two studies.

3.1 The Prime for Life program

Prime for Life is a two-day prevention program based on the “Lifestyle Risk Reduction Model.” The program has been developed by the Prevention Research Institute (PRI) in Lexington, Kentucky (U.S.A.) and is described as a course in risk reduction, and a “(...) program designed to gently but powerfully challenge common beliefs and attitudes that directly contribute to high-risk alcohol and drug use”. ⁷

The PfL program is an example of a type of social alcohol prevention program that has gained popularity in Sweden since the shift from a focus on national alcohol prevention efforts towards the development of preventive measures at the local level. In Sweden, PfL is promoted by PRI as an evidence-based program⁸ and is also endorsed by a number of local municipalities as an evidence-based method “for adults and older teenagers that lowers the risk for all kinds of alcohol and drug problems” (see for example ATAD 2007).

The PfL program aims to teach methods for the reduction of health problems and situation-related problems such as driving while intoxicated, violence and the injuries that can result from alcohol consumption. The goals of the program are:

To help participants to prevent all forms of alcohol- and drug-related problems

To help participants who make high-risk choices to become aware of and accept the need for change in order to prevent future problems.

⁷ www.askpri.org (2012-06-29)
⁸ See for example http://www.primeforlife.se/faq.htm (2012-06-26)
The course promotes the practice of “five steps to risk reduction” to reduce the risk for all kinds of alcohol-related problems. The five steps aim to help participants to (PRI 2003):

1. Recognize signs of biological risks for alcohol-related problems
2. Know the guidelines for “low-risk choices” regarding alcohol consumption
3. Reduce the risks for situational problems
4. Make a decision on whether or not to change alcohol behaviour by identifying which issues are important in such a choice
5. Identify the knowledge and skills necessary to make low-risk choices, for example by choosing activities that offer alternatives to alcohol consumption.

Participants individually assess their guidelines for “appropriate” alcohol consumption on the basis of presentations during the program. These guidelines are either not to drink at all, or to drink a maximum of two units of alcohol per day or not more than three units per setting for those who do not drink every day.

3.1.1 Target group

PRIME For Life can be taught as a prevention program for virtually any group. However, unlike many other drug or alcohol programs, PRIME For Life has been specifically designed for groups that typically make high-risk choices.9

Approximately 80 per cent of the participants in PfL programs in the United States are Driving Under the Influence (DUI) offenders.10 Several U.S. states have treatment programs for DUI-offenders where Prime for Life is included. The program is designed primarily for adult alcohol consumers, and was originally intended to be an efficient therapeutic education for people who make high-risk choices when consuming alcohol.11 The intended principal use of Prime for Life is as a form of secondary or tertiary prevention, and the

---

10 Interview with co-founder the Prevention Research Institute, Ray Daugherty, who has written the Prime for Life curriculum together with Terry O’Brian (2004-12-07).
11 www.askpri.org (2012-06-28)
program makes no claim to immediately have an effect on people with low-risk alcohol consumption.

In Sweden, the main way in which the program has been used is by offering it to broad target groups in which the participants do not necessarily have a high-risk consumption of alcohol. This use of the program can be seen as a form of primary prevention in relation to alcohol problems, i.e. as a program directed at an entire population with the intention of reducing the incidence of alcohol problems (Moskowitz 1989). For example, the program has been used in high schools in Stockholm where, prior to the program, about one-third of the students had claimed to consume alcohol once a month or less (Sjölund & Andréasson 2004).

### 3.2 Evaluations and evidence relating to the Prime for Life program

In the U.S, PfL is included in the Substance Abuse and Mental Health Services Administration’s registry of evidence-based programs and practices (NREPP 2009). The program’s inclusion in the registry is based on two studies of individuals who had been referred to PfL as a state-mandated alcohol and drug education program following involvement in drug-related offenses such as DUI, underage drinking, or drug possession. The first study was a quasi-experimental study among individuals who had received their drug-related sentences and were required to complete a program as a condition for having their driver’s license returned. Outcomes regarding “accuracy of risk estimation” were studied both among participants and in a comparison group that had received a standard intervention (NREPP 2009). The second study was also a quasi-experimental study of participants referred to PfL as a state-mandated program, and measured “intention to drink less or use less drugs” by comparison with a group receiving a standard treatment (Lowenkamp, Latessa & Bechtel 2007). The results were associated with a small effect size in both studies (for an overview of effect sizes, see Chapter 4.4.1).

---

The program is sometimes referred to as evidence-based because it is based on elements and models that are research- and/or evidence-based rather than because the program itself has been proven to be effective. It should be noted that, to my knowledge, only four scientific outcome evaluations of the Prime for Life Program have been conducted in Sweden and all of these have studied the program in a context where it had been implemented as a primary prevention intervention. Two are the quasi-experimental evaluations that I will present in the following chapters, one is a randomized controlled study of the program for high school students in Stockholm and one is a quasi-experimental study of the program conducted among military conscripts. All four studies indicated that the program had a small or trivial effect on alcohol consumption (see Chapters 4 and 5, Sjölund & Andréasson 2004, Hallgren et al. 2009, Leifman 2008).

PfL was accredited by the Swedish Prison and Probation Service in 2008, which means that it is being used throughout the Prison and Probation Service in Sweden. The reason for the accreditation is that the program design is considered likely to deliver desired results, i.e. it is considered to be based on strong theoretical models and evidence-based practices. However, it should be noted that there are as yet no reported outcome evaluation results of the program’s effects on clients in the Swedish Prison and Probation Service.

It is likely that the drive for evidence-based practices (as illustrated in Chapter 2) has given the label “evidence-based” an important marketing value for those who work with designing, producing, promoting, selling and implementing prevention efforts aimed at individuals and groups in Sweden. It could be discussed whether or not it is accurate to describe the PfL program as evidence-based, at least when delivered in the form of a primary prevention effort. Further, I think it is important to understand the PfL program as constituting one of many examples of the type of alcohol prevention initiative that has gained popularity and is being promoted (and discussed!) as an evidence-based practice in an era of individualised alcohol prevention and in the context of a search by local policy makers and professionals for prevention strategies “that work.”

14 Information about the status of the PfL program and of evaluations in the Swedish Prison and Probation service was provided by Martin Larden, Head of Central Coordination in Treatment at the Swedish Prison and Probation Service (2012-06-26). Information about the accreditation of PfL can be found at http://www.kriminalvarden.se/upload/Informationsmaterial/Prime_for_life_1003.pdf (2012-06-26)
3.3 Shared features of my evaluations of Prime for Life

The results that will be presented in Chapters 4 and 5 come from two evaluations that shared the aim of applying quasi-experimental designs to study the impact of the PfL program on risk awareness and alcohol behaviour. Both quantitative and qualitative methods for data collection and analysis has been used in these two studies. In the study at Örebro university statistical analysis of statistical analysis of outcome data from questinnaries to program participants has been combined with qualitative interviews with both participating students and with different stakeholders. The study in the Swedish Armed Forces has been focused on quantitative analysis of outcome data. Both studies have used a program theoretical framework to structure data collection and analysis of both quantitative and qualitative information. In this chapter I will describe the shared components of both studies with regard to data collection, analysis and ethical considerations. Each study and its specifics (such as an implementation study included in the Örebro University evaluation) will be further described in detail in Chapter 4 and Chapter 5 respectively.

3.3.1 Program theory for the Prime for Life Program

All social activities are grounded in some kind of idea or rough draft of how they are supposed to work (Rossi, Lipsey & Freeman 2004). These ideas are rarely documented in detail, but they nonetheless exist in some form of common conception that is shared, for example, by program owners or politicians.

Program theory (also referred to as logical framework, performance framework or theory of action) is a form of theory-driven evaluation perspective that makes use of a theoretical framework to study how an intervention is implemented and how goals and outcomes are produced (Sandberg & Faugert 2012). The scientific assumption guiding theory-driven evaluations is that it is insufficient to judge a social intervention based upon its effectiveness or benefits (Chen 1994). By showing the theoretical assumptions that underlie an
intervention and by learning how the intervention reaches its goals, this evaluation approach can “open up the black box” of social interventions (Harachi et al. 1999) and create an understanding of how and why outcomes are (or are not) produced by a given intervention (Morén & Blom 2007).

Reconstructing the program theory that underlies a given intervention involves illustrating the “path” from a given intervention to its outcomes and the changes it is intended to produce. The theory can be reviewed and assessed on the basis of knowledge made available by research in the field of alcohol prevention for example. In addition, empirical questions can also be identified. These might, for example, be questions about the extent to which the implementation of the activity is in accordance with the program theory, whether the assumptions made by the program theory are the most relevant and whether the activity has created the required conditions to reach its goals.

For the two evaluations presented here, I have derived the basic program theory for the PfL program from the PfL curriculum (PRI 2003) and by participating as an observer in a PfL training course held for staff at the Swedish Prison and Probation Service in 2004. Once the program theory had been outlined (see Figure 1, below), it was validated by means of a personal discussion with Ray Daugherty, who is the president of PRI and the co-author of the Prime for Life Program (2004-12-04).

**Figure 1: Program theory used in the evaluation of Prime for Life**
The program theory consists of three sections (1-3 in Figure 1, above), which can be seen as anticipated steps and outcomes in a process intended to change alcohol use among participants in the PfL program. In short, the program theory illustrates that if (1) the implementation of the program is conducted in a satisfying way, it is assumed that this will produce (2) conditions for the participants to receive knowledge that leads to (3) an increase in risk awareness. The increased risk awareness is then supposed to be a prerequisite for changed behaviour in relation to alcohol consumption (4a, 4b). I will describe each step and how I have approached it in the following:

1. Implementation of the PFL program means that the program has been executed, and that it has been executed under favourable conditions. This part of the program theory is based on the assumption that the context in which the program is implemented is important for the program’s outcomes. It is assumed that motivation to participate in an intervention does not happen automatically. A well-executed program requires that different stakeholders recognize the program as meaningful (Lilja, Giota & Larsson 2004, Sandberg & Faugert 2012). Implementation has been studied from a participant perspective, i.e. with a focus on the participants’ satisfaction with the PfL program. In the study at Örebro University, the implementation of the program has been more thoroughly studied by means of interviews with different stakeholders.

2. That participants receive knowledge and accept it basically means that there is a process of learning in the context of which the participants feel that the program content is trustworthy, and that they obtain new knowledge by participating. Information that is not seen as trustworthy and useful is here assumed to be of little value for producing knowledge that could lead to a change in behaviour among participants. This part of the program theory has been studied by means of questionnaires distributed to participants two weeks after their completion of the PfL program. In the Örebro University evaluation, I have also conducted interviews with students who participated in the program.
3. Received knowledge is assumed to lead to increased knowledge of alcohol-related risks among participants. Without risk awareness it is not likely that the Prime for Life program will lead to a change in alcohol behaviour since the concept of risk is essential to the program’s “change model”. Risk awareness has been measured using a set of questions distributed to both program participants and comparison groups prior to and one year after the PfL program.

The ideal outcome of the PfL program is changed alcohol behaviour, i.e. a reduction in alcohol consumption and/or alcohol-related problems. The intended main use of Prime for Life is as a secondary or tertiary prevention measure, and the program makes no claim to immediately have an effect on people with a low-risk consumption of alcohol; however the program is presented as an alcohol prevention program that can be taught to any group. I have therefore chosen to study outcomes in terms of changes in risk awareness and alcohol behaviour on two different levels:

4. A) Changes in alcohol behaviour among all participants of the PfL program. If the program is effective as a primary preventative effort, then the desired result would be that significant changes occur in the participants’ alcohol behaviour as a result of the program.

B) Changes in alcohol behaviour among participants who prior to the program showed signs of risky alcohol consumption. This would be the most relevant desired outcome, since the program was originally designed as a secondary or tertiary prevention measure.

This simple overview of the PfL program has been used to facilitate a better understanding of the kinds of assumptions that are built into the program and to help structure the data collection, analysis and reporting of evaluation results.
3.3.2 Outcome variables

Both evaluations aimed to investigate the outcomes of the program on risk awareness and alcohol behaviour. For this purpose, multiple variables have been used. These variables were measured both prior to the program (T1) and at a 12-month follow up (T2). I also had a short-term follow up for the PfL program participants two weeks after the program, which focused on evaluating participant satisfaction with the program.

Risk awareness

If the program theory of Prime for Life is correct, increased risk awareness among participants constitutes a necessary first step towards changed alcohol behaviour. “Risk awareness” has been measured in the questionnaires using seven statements derived from the Prime for Life curriculum (PRI 2003).

1. Anyone can develop alcoholism
2. How often alcohol is consumed affects the risk of developing alcoholism
3. How much alcohol is consumed affects the risk of developing alcoholism
4. Having a high alcohol tolerance is a sign of an increased risk of developing alcoholism
5. We are all born with a certain alcohol tolerance that is determined by our genes
6. Every time you drink until you are intoxicated and a little bit more, your tolerance is increased
7. Drinking large amounts of alcohol on a single occasion involves a higher risk for problems than drinking less but more often

The respondents answered by stating their level of agreement with the statements on a five grade scale ranging from “Don’t know” (0 points) to “Agree fully” (4 points). A high level of agreement is in this context to be understood as having opinions about alcohol use that are

---

15 The full set of alcohol-use related questions (in Swedish) can be found in Appendix 2.

16 The statements have been translated from Swedish for purposes of this presentation.
in line with the message of the Prime for Life curriculum (see Question Ec26, Appendix 2). In the evaluations, the participants’ responses to the statements have been combined to form an index. This means that each respondent has been able to score between zero and 28 points on the risk-awareness index.

**Measuring alcohol use and alcohol use disorders**

The evaluation has employed the Swedish version of the Alcohol Use Disorder Identification Test (AUDIT-10). The AUDIT-10 scale consists of ten questions (Appendix 1). Three quantity/frequency questions measure “risky alcohol consumption”, three CAGE questions measure “dependency symptoms” and four questions are about consequences (blackouts, injuries, DUI, physicians’ advice etc.) and are used as a measure of “harmful alcohol consumption” (see Figure 2, below). The scores from the different questions are added together, and an individual respondent’s computed total AUDIT-score can then be used to assess whether or not the individual has “risky” alcohol consumption (Babor et al. 2001).

<table>
<thead>
<tr>
<th>Domains</th>
<th>Question Number</th>
<th>Item Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazardous</td>
<td>1</td>
<td>Frequency of drinking</td>
</tr>
<tr>
<td>Alcohol Use</td>
<td>2</td>
<td>Typical quantity</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Frequency of heavy drinking</td>
</tr>
<tr>
<td>Dependence Symptoms</td>
<td>4</td>
<td>Impaired control over drinking</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Increased salience of drinking</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Morning drinking</td>
</tr>
<tr>
<td>Harmful Alcohol Use</td>
<td>7</td>
<td>Guilt after drinking</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Blackouts</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Alcohol-related injuries</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Others concerned about drinking</td>
</tr>
</tbody>
</table>

*Figure 2: Domains and Item Content of the AUDIT Test (Babor et al. 2010)*
Each of the ten AUDIT questions has a set of responses and each response has a score ranging from 0 to 4. Response scores are added together to produce a total AUDIT-score. In most cases, the total AUDIT-score will reflect the level of risk associated with a given individual’s alcohol use (Babor et al. 2001). There is scientific support for the use of different cut-off scores for men and women (Hradilova Selin 2006). In the analysis, a score of 6 or more points has been used as the cut-off to identify “risk consumers” of alcohol among women. For men, the cut-off score was 8 points.

Alcohol behaviour (use) has been measured using the first three questions of the AUDIT-10 scale, which deal with frequency of alcohol consumption, number of drinks on a typical drinking day and how often the respondent consumes six or more drinks on one occasion (i.e. intensive consumption or “binge drinking”).

**Participant satisfaction**

In both studies, participants were asked to assess the program on a short-term program evaluation form. Overall participant satisfaction will be presented using four characteristics of the program:

1. Overall impression of the program
2. The instructors’ teaching style
3. The instructors’ knowledge
4. The program textbook and its content

Participants graded the program for each aspect on a four-grade scale (very good, good, bad or very bad).
3.3.3 Questionnaire design and testing

The questions were constructed to capture the central aspects of the PfL program and its impact. The Prime for Life curriculum was used to create questions about risk awareness and to get an idea about the theoretical framework and goals of the program. I also consulted Ray Doroughty, president of PRI. Additional support in the process of selecting questions and designing the questionnaires was received from Robin Room, Professor and Director of SoRAD (1999–2006) and Sandra Bullock, Ph.D. in Public Health Science and post-doc researcher at SoRAD at the time of the evaluations. The original (Swedish) version of the questionnaire that was used (at T2 in the evaluation of PfL in the Swedish Armed Forces) can be found in Appendix 2.

Questionnaire content and usability was tested both among strategically selected personnel in the Swedish Armed Forces as well as among fellow alcohol researchers at the Centre for Social Research on Alcohol and Drugs (see Chapter 5).

3.3.4 Statistical analysis

Those individuals who participated in the follow-up study (T2) have been matched individually with their responses in the pre-test (T1). This means that the evaluations only include those individuals who participated at both T1 and T2. The matching has been conducted using slightly different techniques and will be described in the presentation of the results from each evaluation (Chapters 4-5).

The method last measure forward (LMF) has been used in the analysis in both studies. LMF means that if a value on the outcome variable has been missing at T2, the value reported by the respondent at T1 has been used.
Outcome results presented as Effect Sizes

In the original reporting of evaluation results (Sandberg 2006, Sandberg 2007a, Sandberg 2007b), independent t-tests and non-parametric statistics (Wilcoxon and Mann-Whitney) were utilized. In this thesis, I have focused on the outcome of the PfL program in terms of effect sizes since this is considered good practice when presenting empirical research (Wilkinson 1999).

Effects of Prime for Life are based on Cohen’s $d$ on the key outcomes twelve months after the program (T2). Each group’s standard deviation has been weighted by its sample size ($n$) and Cohen’s $d$ has been calculated from analyses of covariance (ANCOVA), adjusting for pre-test differences in outcome and relevant background variables. This statistical approach to accounting for baseline measurements is recommended by the Cochrane Collaboration (Deeks et al. 2008).

When assessing the size of ES, I have used the following criteria (SBU 2001):

- Less than 0.20: Trivial effect
- 0.20–0.50: Small effect
- 0.50–0.80: Moderate effect
- 0.80 or higher: Large effect

It should be noted that even small effect sizes may be of importance within a particular area of study (Schuele & Justice 2006), although smaller effect sizes are generally viewed as questionable in the field of alcohol prevention (Foxcroft 2006). In my studies I have chosen to follow the above criteria, which mean that an ES of 0.50 or higher has been the benchmark for identifying interesting and policy-relevant outcomes.

---

17 The results were not adjusted for pre-test differences in outcome and background variables in connection with the original reporting of results (Sandberg 2006, Sandberg 2007a, Sandberg 2007b). The adjusted calculations of Cohen’s $d$ have not affected the previously reported conclusions regarding the effectiveness of the program.

18 The Cochrane Collaboration is an independent nonprofit organization that was formed to organize medical research information in a systematic way in the interests of evidence-based medicine.
3.3.5 Ethical considerations

In both evaluations, all participants were informed that participation was voluntary and that they could drop out of the evaluation at any time. Measures have also been taken to ensure anonymity while maintaining the ability to individually match pre- and post-test data. For this reason, the questionnaire and data collection designs were formulated to ensure that it would be impossible for me to identify individuals in the data sets. (The data collection procedure will be described for each study in Chapters 4 and 5.)

The evaluation plan for the evaluation study at Örebro University was sent to the regional ethical review board in Stockholm. The committee decided that the research project did not require to be ethically reviewed since the research subjects had given their consent to participate. Since the study conducted within the Swedish Armed Forces employed a similar design and data collection procedure, its evaluation plan was not submitted to the review board.

Outcome results from the evaluation of Prime for Life within the Swedish Armed Forces have been analysed for both men and women. At Örebro University, I have decided not to present results based on gender because it might then be possible to identify individual students in the evaluation. Because of the sensitivity of the topic at hand (alcohol consumption and alcohol-related problems) and the small study population, I have also made the decision not to reveal the identity of interview persons in the reporting of the results from the evaluation at Örebro University.
4. Study 1: Evaluation of the Prime for Life Program at Örebro University

This chapter will present the evaluation and the results of the PfL program at the Department of Hospitality, Culinary Arts and Meal Science at Örebro University (Sandberg 2006, Sandberg 2007a).

4.1 Introduction

In 2004 the National Alcohol Committee\textsuperscript{19} contacted Örebro University with the intention to pilot the Prime for Life Program in a university setting. The Prime for Life Program was implemented at the Department of Hospitality, Culinary Arts and Meal Science at Örebro University in the spring of 2005. The department is located on its own campus about 90 kilometres north of the main Örebro University campus. In the spring of 2005, there were 223 students registered at the department.

In the fall of 2004, the PfL program was offered to staff at the department as a way of obtaining approval for the implementation of the program among the department’s students in the spring of 2005. The idea was that teachers and staff, after participating in the program themselves, would be better prepared to tell the students about the program and to encourage them to participate, and the program was then implemented as a voluntary course for all students at the department.

The evaluation was initiated in the summer of 2004. Sixty-three students participated in the PfL program in May 2005, and 44 of these (70 \%) were included in the evaluation.

The evaluation was commissioned and financed by the National Alcohol Committee.

\textsuperscript{19} The National Alcohol Committee was commissioned by the government to implement the National Alcohol Policy (prop. 2000/01:20) during the period 2001–2007. The goal of the National Alcohol Committee was to coordinate and promote initiatives to reduce alcohol-related harms in Sweden.
4.1.1 Study objectives and research questions

The aim of the evaluation was to investigate the impact of the PfL program on the participants’ alcohol behaviour. In addition, the Alcohol Committee wanted to study whether an intervention such as Prime for life could be useful for alcohol prevention initiatives at universities in Sweden.

The study had two outcome-oriented research questions:

1. Are there any changes in participants’ risk awareness?
2. Are there any changes in participants’ alcohol behaviour?

In addition, the Örebro evaluation included four specific program- and implementation-related questions:

1. Does the program suit the target group?
2. How is the program received by the participants?
3. Does the program lead to any changes in alcohol culture?
4. What possibilities and problems are associated with implementing a program like Prime for Life in a university setting?

4.2 Method

Qualitative data were collected in the form of semi-structured in-depth interviews with instructors and administrative staff at the Department of Hospitality, Culinary Arts and Meal Science and in structured interviews with students who had participated in the PfL program. In addition, regular meetings were held throughout the evaluation period with PRI representatives, a representative of the Student Health Services at Örebro University and a representative from the National Alcohol Committee.

Quantitative outcome data were collected by means of questionnaires among students who had participated in the program. Questionnaires were administered two weeks prior to the program, two weeks after the program and 12 months after the program. Comparative data were collected from a comparison group of students at Örebro University who did not
participate in the Prime for Life program. The outcome study was designed as a quasi-experimental study since it had already been decided when the evaluation was initiated by the Alcohol Committee and Örebro University to offer the Prime for Life Program as a voluntary program to all students at the department.

4.2.1 Interviews with teachers and staff at the department

Interviews with instructors and staff at the Department of Hospitality, Culinary Arts and Meal Science were conducted in order to obtain information about the prerequisites for the implementation of Prime for Life. The interviews were also intended to gather information about the staff’s experience of, and opinions about the PfL program.

In April 2005, a 120 minute long group interview was conducted with instructors and staff at the department. Six individuals from the department participated in the interview, which was conducted in the form of a semi-structured in-depth interview. The interview was recorded and subsequently transcribed. All of the participants in the group interview had participated in the PfL program offered to departmental staff in the fall of 2004. The interview focused on the staff’s experience of the program, and their thoughts on the program being implemented among students at the department. The group interview encouraged the participants to discuss their experiences and thoughts with one another. To complement the group interview, telephone interviews were conducted between August and October 2005 with six strategically chosen departmental representatives who had different interests in relation to the implementation of PfL at the department.

4.2.2 Interviews with students

In the fall of 2005, about six months after the PfL program, structured telephone interviews were conducted with 30 students who had participated in the program at the department.

---

Tove Sohlberg at SoRAD assisted in the group interview.
The interviews included ten open-ended questions covering the students’ opinions about the PFL program, their image of the alcohol culture at the department and what they thought about possible outcomes of the program.21

4.2.3 Questionnaires

The data for the outcome study were collected by means of a web-survey. This meant that questionnaires were answered via a web page, which the students obtained access to by means of a personal link that had been sent to participants in an e-mail. To answer the questionnaire, the students had to use the link and log in with an individual password that was included in the e-mail along with the link. After answering the questionnaire online, each respondent’s answers were saved together with an identification code in a secure database hosted by the statistical consultant company Statisticon in Uppsala.

Because there was no information available about which students would choose to participate in the Prime for Life Program, and since the Alcohol Committee wanted to leave the opportunity open for students to make a late decision to participate, the first questionnaire (T1) was sent out to all (223) students registered at the department. After the program had been implemented, those students who had not participated in the program were removed from the dataset by Statisticon.

The reason for performing a web-survey was that the management at the Örebro University noted that e-mail was widely used to communicate with students about courses, examinations etc. University and departmental information is distributed to the student’s university account, an e-mail account that the students are given when they register at the university. From a research perspective, this situation provided an opportunity to try and assess a “new” method of collecting data. Postal questionnaires were distributed to students who did not reply to the web-survey.22

21 The telephone interviews were conducted by Alpha Sow at SoRAD.
22 Of the data employed in the outcome evaluation, approximately 85 percent came from the web-survey and 15 percent from postal questionnaires.
Data for the pre-test were collected in April–May 2005 and the follow-up data in April–May 2006.

**Experiment group**

The experiment group comprises the 63 students who completed the PfL program. Of these, 44 (70 %) participated in both the pre-test (T1) and the post-test (T2) 12 months after the program.

**Comparison group**

For practical reasons, the comparison group was drawn from other disciplines at Örebro University. Using information (sex and age distribution) available from the university, students from the Department of Finance, Economy and Informatics and the Department of Natural Sciences were chosen as control group. The advantage of this procedure may be that all participants in the evaluation are students who are exposed to the same university policies and action plans regarding alcohol and who share the same access to Student Health facilities and the like. However, unlike students at the Department of Hospitality, Culinary Arts and Meal Science, the students in the comparison group live in the city, and have a larger campus area. This is an important shortcoming of the evaluation, since access to restaurants and bars can be an important factor when it comes to alcohol consumption among students (see also Chapter 8).

Based on a “best case scenario” that about 200 students would choose to participate in the Prime for Life program at the Department of Hospitality, Culinary Arts and Meal Science, an oversized random sample of 1,200 students at the two departments was drawn for the comparison group. Of these, 359 students (29 %) who participated in both the pre-test (T1) and post-test (T2) are included in the evaluation study.
4.3 Quality of outcome data

There is most likely a selection bias in the study, since the program was voluntary and no randomization was conducted in relation to the participants. Since I had collected data for the pre-test prior to the implementation of Prime for Life at the Department of Hospitality, Culinary Arts and Meal Science based on a sampling frame comprising all registered students at the department, I have had access to data from some students who chose to complete the pre-test but not to participate in the program. This has allowed me to make some comparisons between the study population, i.e. PfL participants, and other students at the department. It should be noted, however, that the only data available for this comparison are the data from the web-survey, which means that Table 1 below compares only a subset of 36 of the 44 students who participated in the prime for Life program with 34 students from the department who completed the questionnaire but decided not to participate in the program.

Table 1: Comparison of outcome variables at T1 between experiment group (n=36) and students from the department who did not participate in the PfL-program. (Web-survey data)

<table>
<thead>
<tr>
<th></th>
<th>Experiment group (n)</th>
<th>Non-participants (n)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk awareness</td>
<td>20.92 (36)</td>
<td>20.03 (31)</td>
<td>-</td>
</tr>
<tr>
<td>Drinks on a typical day drinking alcohol</td>
<td>2.33 (36)</td>
<td>2.34 (33)</td>
<td>-</td>
</tr>
<tr>
<td>How often alcohol consumption/month</td>
<td>2.97 (36)</td>
<td>3.03 (34)</td>
<td>-</td>
</tr>
<tr>
<td>Frequency binge drinking</td>
<td>2.34 (36)</td>
<td>2.31 (31)</td>
<td>-</td>
</tr>
</tbody>
</table>

As Table 1 (above) shows, the limited data available do not indicate any selection bias in terms of differences between PfL participants and other students at the department. However, it is possible that the students who answered the questionnaire at T1 but then decided not to participate in the program had initially been interested in the program and had intended to participate. The “non-participants” in Table 1 might therefore be more similar to those who participated in the program than students at the department who had
no interest in participating in the program. Therefore, the possible problem of selection bias should still be taken into consideration – this issue was also discussed by some of the students in the interviews, see below (Chapter 4.5.5).

In the following tables (2–3), background and outcome variables are compared between the experiment and the comparison group at the time of the pre-test (T1).

**Table 2: Comparison of background data between experiment (N=44) and comparison group (N=354), Örebro University. Percent.**

<table>
<thead>
<tr>
<th></th>
<th>Experiment group</th>
<th>Comparison group</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>57</td>
<td>59</td>
<td>-</td>
</tr>
<tr>
<td>Men</td>
<td>43</td>
<td>41</td>
<td>-</td>
</tr>
<tr>
<td>Born before 1980</td>
<td>34</td>
<td>60</td>
<td>***</td>
</tr>
<tr>
<td>Born after 1980</td>
<td>66</td>
<td>40</td>
<td>***</td>
</tr>
</tbody>
</table>

***p<0.01

The gender distribution is similar in the experiment and comparison group, but at the time of the pre-test, the experiment group included a statistically significant larger proportion of students under 25 years of age.

**Table 3: Comparison between experiment (N=44) and comparison group (N=359) at T1, Örebro University. Mean values.**

<table>
<thead>
<tr>
<th></th>
<th>Experiment group</th>
<th>Comparison group</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk awareness</td>
<td>20.73</td>
<td>17.75</td>
<td>***</td>
</tr>
<tr>
<td>Drinks on a typical day drinking alcohol</td>
<td>2.25</td>
<td>2.39</td>
<td>-</td>
</tr>
<tr>
<td>How often alcohol consumption/month</td>
<td>3.00</td>
<td>2.90</td>
<td>-</td>
</tr>
<tr>
<td>Frequency binge drinking</td>
<td>2.30</td>
<td>2.16</td>
<td>-</td>
</tr>
<tr>
<td>Total AUDIT-score</td>
<td>6.83</td>
<td>6.52</td>
<td>-</td>
</tr>
</tbody>
</table>

***p<0.01

The experiment group had a statistically significant higher mean risk awareness by comparison with the control group. This difference is probably a result of how the
comparison group was chosen for the study; many students at the department have participated in a course called Responsible Beverage Service, which is a requirement for the students to be able to obtain their degree from the department. Responsible Beverage Service focuses on awareness about problems associated with alcohol consumption (FHI 2009b) and therefore is in some ways similar to the risk awareness concept taught in the PfL program. It is likely, however, that the prevalence of such prior training was limited in the comparison group. It should be noted that despite this initial difference in risk awareness, there were no statistically significant differences between the two groups regarding the alcohol use variables.

4.4 Main outcome results of the evaluation

In the following, the outcomes of the Prime for Life Program at Örebro University on risk awareness and alcohol behaviour will be presented. Please note that the data have not been analysed by gender (men and women) due to the small size of the experiment group. Results have been adjusted for pre-test differences for each outcome variable and age.
**Table 4. Effects of Prime for Life (PFL) on key outcomes 12 months after the program (T2). Results from analyses of covariance (ANCOVA), adjusting for pre-test differences in outcome and age (full sample, 95% confidence intervals in parentheses).**

<table>
<thead>
<tr>
<th>Variable\ outcome</th>
<th>Risk awareness</th>
<th>No. of drinks on a typical day drinking</th>
<th>How often alcohol consumption</th>
<th>Frequency binge drinking</th>
<th>Total AUDIT score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
</tr>
<tr>
<td>PFL</td>
<td>4.03</td>
<td>0.91</td>
<td>-0.18</td>
<td>-0.14</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of observations</th>
<th>n_{total}=389</th>
<th>n_{PFL}=44</th>
<th>n_{control}=345</th>
<th>n_{total}=394</th>
<th>n_{PFL}=44</th>
<th>n_{control}=350</th>
<th>n_{total}=377</th>
<th>n_{PFL}=43</th>
<th>n_{control}=333</th>
<th>n_{total}=361</th>
<th>n_{PFL}=42</th>
<th>n_{control}=319</th>
</tr>
</thead>
</table>

SMD = Standardized mean difference (Cohen’s d).

The increase in risk awareness among PFL participants was large (SMD=0.9). There were trivial or small effects on alcohol behaviour.

**4.4.1 Outcomes for risk consumers**

The following analysis includes female students with an AUDIT-score of 6 points or more and male students with an AUDIT-score of 8 points or more.
TABLE 5. EFFECTS OF PRIME FOR LIFE (PfL) ON KEY OUTCOMES 12 MONTHS AFTER THE PROGRAM (T2). RISK CONSUMERS. RESULTS FROM ANALYSES OF COVARIANCE (ANCOVA), ADJUSTING FOR PRE-TEST DIFFERENCES IN OUTCOME AND AGE (FULL SAMPLE, 95% CONFIDENCE INTERVALS IN PARENTHESES).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Risk awareness</th>
<th>No. of drinks on a typical day drinking</th>
<th>How often alcohol consumption</th>
<th>Frequency binge drinking</th>
<th>Total AUDIT score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
</tr>
<tr>
<td>PfL</td>
<td>3.97</td>
<td>0.85</td>
<td>-0.36</td>
<td>-0.33</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>[2.15 – 5.73]</td>
<td>[0.49 – 1.11]</td>
<td>[-0.80 – 0.08]</td>
<td>[-0.57 – 0.09]</td>
<td>[-0.04 – 0.16]</td>
</tr>
<tr>
<td></td>
<td>-0.85</td>
<td>0.33</td>
<td>-0.08</td>
<td>-0.14</td>
<td>-0.25</td>
</tr>
<tr>
<td></td>
<td>[-0.57 – 0.08]</td>
<td>[-0.38 – 0.10]</td>
<td>[-0.04 – 0.16]</td>
<td>[-0.51 – 0.01]</td>
<td>[-0.51 – 0.01]</td>
</tr>
<tr>
<td></td>
<td>-0.36</td>
<td>0.14</td>
<td>-0.08</td>
<td>-0.40</td>
<td>-2.08</td>
</tr>
<tr>
<td></td>
<td>[-0.80 – 0.08]</td>
<td>[-0.38 – 0.10]</td>
<td>[-0.04 – 0.16]</td>
<td>[-0.64 – 0.16]</td>
<td>[-3.52 – -0.64]</td>
</tr>
<tr>
<td></td>
<td>-0.33</td>
<td>0.25</td>
<td>-0.25</td>
<td>-0.53</td>
<td>-0.53</td>
</tr>
<tr>
<td></td>
<td>[-0.57 – 0.09]</td>
<td>[-0.38 – 0.10]</td>
<td>[-0.51 – 0.01]</td>
<td>[-0.64 – 0.16]</td>
<td>[-0.69 – -0.37]</td>
</tr>
</tbody>
</table>

SMD=Standardized mean difference (Cohen’s d).

For risk consumers of alcohol, there was a large effect of the PfL program on risk awareness. There was a decrease in the number of drinks on a typical drinking day, in how often alcohol was consumed and also in the frequency of binge drinking. In terms of effect sizes, however, the changes were trivial or small. For risk consumers, the program had a moderate effect on the mean AUDIT-score.
4.4.2 Comments regarding the outcome results

The students who participated in the Prime for Life Program significantly increased their risk awareness, which seems reasonable since this is what the program sets out to teach participants about.

For participants in general, the evaluation could only identify trivial or small effects of the program on alcohol behaviour (i.e. frequency of consumption, number of drinks consumed and frequency of binge drinking). Since the Prime for Life program was primarily intended as an intervention for risk consumers, this result is perhaps not surprising.

For students who could be identified as risk consumers of alcohol, i.e. those who according to the program theory of PfL were most likely to change their alcohol behaviour as a result of participating in the program, there were also trivial or small effects of the program on alcohol behaviour – with the exception of the total AUDIT-score where there was a moderate effect of the program.

**Table 6: Proportion of the total AUDIT change for different domains of the AUDIT Test. Risk consumers, Örebro University. Percent.**

<table>
<thead>
<tr>
<th>Question No.</th>
<th>Domain</th>
<th>% of total change (total points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hazardous use</td>
<td>38 (0.98)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Dependency symptoms</td>
<td>12 (0.31)</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Harmful Alcohol use</td>
<td>51 (1.32)</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This simple analysis indicates that the largest proportion of the change in mean AUDIT-scores for risk consumers who participated in the Prime for Life program was due to a change in their scores relating to the Harmful Alcohol use domain of the AUDIT test. Fifty percent of the difference in the mean AUDIT-score score can be explained by changes in reported harmful alcohol use; i.e. frequency of guilt or memory loss due to alcohol consumption and whether drinking has caused any injuries or led to a doctor or family member suggesting that the person should cut down on their drinking during the past 12 months. Thirty-eight per cent of the difference in the mean AUDIT-score for risk consumers is related to changes in hazardous use, i.e. frequency of alcohol consumption, number of drinks per drinking occasion and frequency of binge drinking (where a trivial to small program effect could be identified on the outcome variables).

4.5 Results regarding the implementation of the Prime for Life Program

In the following, implementation refers to the study of what is happening and why in the design, implementation, administrative operation, services, and outcomes of the PfL program.

The Alcohol Committee chose, as a result of established contacts and a statement of interest from Örebro University, to conduct a pilot testing of the Prime for Life program at the Department of Hospitality, Culinary Arts and Meal Science. In May 2005 the program was implemented at the department. The following results are from the implementation study conducted in parallel with the outcome study at the Department of Hospitality, Culinary Arts and Meal Science. All interview excerpts have been translated from Swedish by the author.

4.5.1. Department representatives’ experience of Prime for Life

When the evaluation was initiated in the summer of 2004, Alcohol Committee officials described both the Head of Department and the Vice-Chancellor of Örebro University as being “very positive” to the idea of implementing PfL at the Department of Hospitality,
Culinary Arts and Meal Science. By September of the same year, according to the Alcohol Committee, the intervention had been approved by the department.

In November 2004, PRI offered a PfL program for the staff at the department. Less than half of the 20 teaching and other staff at the department participated. PRI instructors reported disappointment following the implementation of the program among departmental staff, since they had felt a resistance towards the program on the part of the staff and said that the staff had behaved nonchalantly towards them during the program.

A second course date was scheduled for January 2005. Instructors and other staff at the department were invited, together with the representatives of the student organization. Only three persons signed up for the program’s second date and it was cancelled.

In the interviews with teachers and other staff from the department, the majority of interviewees expressed disappointment with the Prime for Life program. Examples from two interviewees’ reflections about the program are illustrated below:

(1)

(...) it was kind of like that style that there was in early school-years... push that in, do not reflect, do not think, just learn. So, I thought the program was bad, actually.

(2)

(...) they asked simple questions and limited our opportunities to reflect ourselves. It was embarrassing. You can ask kids in that way, but we are adults. The dialogue with the participants was not good enough.

In the group interview there seemed to be an overall consensus that the staff did not believe that the PfL program would suit their students. One teacher at the department expressed this in the following way:

(...) the students will not come back after the first day of teaching. (---) If they sit there and receive the same information as I did, they will never come back.
Some of the interviewees meant that the type of educational program that PfL involves places substantial demands on the way it is delivered and several indicated that they did not think the teaching style employed by the PfL program instructors was very good; they were also certain that the students would share their opinion about this.

Some members of departmental staff provided a somewhat more nuanced description of the program. They were of the view that there was a communication problem between the Prevention Research Institute and the staff. They felt that the staff had not been properly informed about the content of the program, and that they were therefore quite suspicious about the PfL program and the program instructors. The Department of Hospitality, Culinary Arts and Meal Science was described by interviewees as being unique and not comparable with other university departments. The staff therefore came to the program with specific ideas about what they needed and what they wanted to discuss. Several interviewees also pointed out that the decision to implement the program had been made by the Head of Department without first discussing it with them.

For the teachers at the department, the PfL program became a burden, since they had to reschedule classes to make room for something that they did not feel that they had initiated. Several of the teachers who had participated in the PfL program at the department did not think that the program was appropriate for their students.

Instructors at the department felt that it was a problem that the implementation of the PfL program came as a “top-down” directive and that they had been forced to adjust their course schedules to fit in with the program. During the group interview, it was discussed whether or not these perceived “difficulties” were because they personally did not like the Prime for Life program. Below are three excerpts from this discussion:
(1)
- (...) if the teaching would have been better, we would not have had so many discussions...
- Then our attitude would have been different
- Yes... then it would not have been so difficult
- Probably not...

(2)
We are continuously evaluating ourselves and we want to be the best in the business and make sure that what we deliver is really top notch... then it felt like, well this PfL thing does not really meet our standards.

(3)
It was difficult to motivate... that other elements in the course schedule were removed for this... when I know what it is...

A follow-up question raised in the context of this discussion was that of whether the staff felt that their own experience of the Prime for Life program might have affected how they then presented the program to their own students. The staff made it clear that they had not given the students a negative description of the program, but rather that they had made an effort to keep a neutral tone when talking about it with their students.

Following the failed attempt to implement the Prime for Life program among the staff at the department, extensive discussions were held between the Alcohol Committee, the department, the student organization, the Prevention Research Institute and the evaluator about strategies to ensure that the program could be successfully implemented among the students. As a consequence, a representative of the Alcohol Committee and myself held an information meeting for students at the department in February 2005, and it was decided to include complimentary lunches in the Prime for Life program for participating students. At the recommendation of the student organization, which thought that the Prime for Life program could provide important additional value to the students’ education and could actually be a merit when applying for a job, it was also decided that participating students would receive a certificate showing that they had participated in the program.
4.5.2 Short-term program evaluation results and student experiences of the program

In the interviews six months after the PfL program, a majority of the students claimed that they had participated in the program simply out of curiosity or a specific interest in alcohol questions (18 of 27 students). Among the other most frequent reasons given for having participated was that the program was included in their course schedule or that the student had thought that participation was compulsory (5 students). Most students (14) could not describe any specific expectations of the program. However, several (11) were pleasantly surprised by the content and high quality of the program.

In the questionnaire administered to participants two weeks after the Prime for Life program, the students assessed the program on the basis of four different factors.

**Table 7: Student grading of Prime for Life. Percent. (N=40), Örebro University**

<table>
<thead>
<tr>
<th></th>
<th>Overall impression of the program</th>
<th>The instructors’ teaching style</th>
<th>The instructors’ knowledge</th>
<th>The program textbook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
<td>17</td>
<td>29</td>
<td>71</td>
<td>29</td>
</tr>
<tr>
<td>Good</td>
<td>73</td>
<td>63</td>
<td>27</td>
<td>61</td>
</tr>
<tr>
<td>Bad</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Very bad</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Don’t know/non-response</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

As a simple course evaluation, the results indicate that the students had a very positive experience of the PfL program. Further, 98 percent of the students in the program found the content of the program trustworthy or very trustworthy, and the vast majority of the
students (88 %) claimed that it was important or very important to have programs like PfL at the university (not shown in table 7. See Sandberg 2006).

Because of the strong reaction from the staff at the department about how the program was delivered by the instructors from the Prevention Research Institute, in the interviews conducted six months after the program the students were specifically asked to compare the teaching style of PfL with the classes given at the department. Three examples of student answers:

(1)

It [the Prime for Life program] was very cutting edge in its delivery. Sometimes courses are not so good, but this was a good presentation. The quality in the teaching at the department is very inconsistent.

(2)

We usually have teachers that just talk and who know nothing about pedagogics. Celebrity chefs come here, people from the business. Sometimes they just have a bunch of slides that they read off. Prime for Life was much better. The instructors were experienced and they knew what they were talking about.

(3)

It varies a lot in our courses. It was positive that Prime for Life was thought through, as that is the missing link sometimes in our education. We had a small group, room for discussions and plenty of time. That was good for us.

The disparity between the teachers’ and the students’ experience of the PfL program is evident.
4.5.3 Alcohol culture at the department

I tried to understand what the alcohol culture (i.e. the customs and practices associated with the consumption of alcoholic beverages) was like at the Department of Hospitality, Culinary Arts and Meal Science. The study environment can be described as unique in the sense that, for example, it could be argued that the education provided is promoting alcohol use. This is because alcohol is part of the program, which provides sommelier (wine professional) training programs as well as courses focusing on wine and spirits. On the other hand, each student has to participate in a course involving responsible alcohol use in order to get their degree at the department. This course involves knowledge about alcohol policy, the dangers of alcohol and the “responsible handling of alcohol” in a restaurant setting. The course provides general information about the medical and social consequences of alcohol use, as well as legal regulations regarding the serving of alcohol in restaurants.

During the course of the evaluation, I experienced a sense of confusion about what the status of alcohol was at the department. While visiting the department and communicating with students as well as department and university representatives, I made observations that indicated an ambiguous attitude towards alcohol. For example, there were some indications that that there was perhaps (or had been) alcohol-related problems at the department, but there was no one who wanted or was able to verify this information for me. The question about alcohol seemed to be a sensitive issue at the department and I was able only to make occasional observations rather than systematically collecting information about the alcohol culture. Consequently, I was not able to come to any clear conclusions about the alcohol culture at the department in the evaluation (which made it difficult to answer the question of whether or not the PfL program had resulted in any changes to the alcohol culture at the place where it was implemented, see below). However, I will try to provide a few examples that illustrate the contradictory images of the alcohol culture that I obtained from various data sources:
The unique curriculum of the department involves the consumption of alcohol during class hours. A person in the group interview conducted with teachers and staff commented upon this by saying the following:

Here, one has the possibility during class hours to test a lot of alcohol... and we recommend all students to spit it out... but of course, a lot is being swallowed. We have not conducted any tests on the students... to see if they can drive a car afterwards... but... well, they drink a lot, so they learn how to drink, but they do learn how to drink responsibly.

The staff at the department who were interviewed did not think that alcohol consumption among students at their department was higher than in other student environments. Rather, the staff made a point of noting that their students were significantly different from other students since they were being trained in how to deal with alcohol in “the right way”. The students “(...) talk about food and drinks in combination, and they talk more about quality than quantity” as one departmental teacher explained it.

Several of the interviewees representing department teachers and staff described the department’s alcohol policy as being “extremely strict”. The department had an alcohol policy long before the university itself formulated its own alcohol policy. According to the participants in the group interview, this policy had been formulated by the students at the department themselves. According to them, most students at the department were well aware of the alcohol policy and its content. However, data from the questionnaire among students who participated in the Prime for Life program (N=44) indicated that only one student claimed that he or she had very good knowledge about the department’s alcohol policy (Sandberg 2006), while almost half (48 %) of the students claimed to have no knowledge about the policy at all.23

The two most common (and also most contradictory) types of answer from the students about the alcohol environment at the department indicated that there was either a problematic alcohol consumption among the students at the department, or that the students at the department were very cautious about alcohol and preferred smaller

23 At T2 (n=41) the level of students saying they had very good knowledge about the policy had increased to 5 percent, while 32 percent claimed that they had no awareness of the alcohol policy at all (Sandberg 2006)
amounts of more exclusive types of alcohol than students at other places. The following four quotes illustrate this:

(1)
We are not many students here – everybody knows everybody. There is always a bunch of students that drink a lot on Wednesdays and Saturdays. There are more parties with exclusive wines. Many students already have bad alcohol habits [from working in bars and restaurants] before they start at the department.

(2)
Restaurant people in general seem to drink a lot. Many people in the business have a high tolerance level. They know a lot about wines, but as students they drink the cheapest wines.

(3)
At Grythyttan [the department] we have more exclusive alcohol behaviour. We spend more money on more exclusive wines.

(4)
We drink to learn. At other places, people drink to get intoxicated. We do not drink just anything. We drink more quality than quantity.

4.5.4 Did the program produce any changes in alcohol culture?

It was impossible to provide a clear answer to this question, partly because of the problems described above in relation to developing a clear picture of the alcohol culture at the department. In addition, the PfL program itself is an individual program that targets the participants’ own relationship to alcohol, rather than that of a group.

In the interviews, ten students claimed that they thought that Prime for Life could change the alcohol culture of the student environment, while just as many thought that the chances of the program having an impact on the alcohol culture at the university were limited. A few students indicated in the interviews that those who would have benefited most from
participating in the program did not show up and that more students would have needed to participate for the program to have an impact at the departmental level.

Those who came [to the PfL training] are those with fewer problems. Those who have problems with their alcohol consumption did not go there.

If this is true, it seems important to consider what would have been needed to produce a higher participation rate. Mandatory participation, for example, might have been one way of ensuring that more risk consumers of alcohol at the department would participate in the program. Also, a more positive (rather than a “neutral”) attitude towards the program among staff at the department might have encouraged more students to participate.24

The evaluation concluded that it was unlikely that the PfL program had contributed to any changes in the alcohol culture, and that for the program to reach a “critical mass” as part of an effort to change the drinking environment where it is implemented, more students need to participate and the program needs to obtain the approval of all stakeholders and should preferably also be linked to some form of systematic and long-term alcohol prevention policy (Sandberg 2006).

4.5.5 The possibility of implementing the program in a university setting

Sixty-three of about 220 students at the Department of Hospitality, Culinary Arts and Meal Science participated in the Prime for Life Program. This may be regarded as a relatively good participation rate given the fact that participation was voluntary and required a two-day commitment on the part of the students, and it is possible that many students could not attend the program simply because they had to prioritise other things. However, the neutral stance that the departmental teachers claimed to have adopted when they talked about the program with their students may have made it appear that the department was not overly...

24 The issue of selection bias has also been addressed in Chapter 4.3.
interested in encouraging the students to participate in the program. Nevertheless, the majority of students participating in the program rated it very highly.

My overall conclusion about program implementation is that a closer cooperation between different stakeholders would have been required to improve and enhance the level of acceptance for and the implementation of the PfL program. And if the program is implemented as a mandatory part of a university education, it may be of value at the group level to create shared ideas about alcohol consumption and problems, which might help to produce approval for alcohol policies and strategies to deal with alcohol-related issues (see also Chapter 7.2).

This chapter describes the outcome evaluation of the PfL program in the Swedish Armed Forces (Sandberg 2007b).

5.1 Introduction

In 1999, the Supreme Commander decided to implement the PfL program at all levels within the Swedish Armed Forces. In 2004, the Centre for Social Research on Alcohol and Drugs (SoRAD) at Stockholm University was asked by the Swedish Armed Forces to evaluate the program. At the time of the evaluation, the Swedish Armed Forces had about 20,000 civilian and military employees at eighteen regiments geographically distributed throughout the country. The evaluation was initiated in 2004, and data for the evaluation were collected between January 2005 and February 2007. The study was commissioned by the Swedish Armed Forces and financed by The Development Council for the Government Sector.

5.1.1 Objectives and research questions

The purpose of the evaluation study was to assess whether the Prime for Life program had any impact on the alcohol behaviour of the armed forces personnel.

The two main research questions for the evaluation were outcome-oriented:

1. Are there any changes in participants’ risk awareness?
2. Are there any changes in participants’ alcohol behaviour?

5.2 Method

The evaluation was conducted in the form of a quasi-experimental study using questionnaires distributed to participants in the Prime for Life program at six regiments two weeks prior to their participation in the program (T1), two weeks after the program and 12 months after the program (T2). Personnel from ten regiments that did not implement the Prime for Life program during the evaluation served as comparison group.
Different levels of motivation and different capabilities in relation to program implementation were expected across different military regiments, and the alcohol- and drug coordinators at military headquarters identified six regiments that were motivated and prepared to implement the PfL program. The goal was to give the program to 1,000 staff members at these regiments over the course of a one-year period. The comparison group consisted of ten regiments that were selected on the basis of geographical position and branch of the armed forces to match the experiment regiments as far as possible. The goal was to obtain the participation of a total of 3,000 employees at the comparison regiments.

5.2.1 Questionnaires

For practical (e.g. financial and administrative) reasons, the Swedish Armed Forces assumed responsibility for collecting the evaluation data. Using staff consultants employed at each regiment, the questionnaires were administered to both military and civilian personnel. To ensure that targets on data collection would be obtained, and that the questionnaires were administered in a similar way at all regiments, detailed written instructions were distributed to the regiments to ensure that respondents were informed both orally and in writing (in the questionnaire) about the fact that their participation in the study was voluntary. Participants were guaranteed anonymity throughout the evaluation and were also informed of their option to withdraw from the evaluation at any time. In addition, evaluation reports were regularly collected from the regiments (see below).

As far as possible, the survey was performed with groups of personnel being scheduled to fill out the questionnaire at the same time. This method was chosen over sending out questionnaires to each of the individual respondents at the regiments, since it allowed for the staff consultant to provide all participants with the same information about the purpose of the study and also to help with questions and to forward any questions or comments to the evaluator. The questionnaire was completed individually and anonymously by each employee. After completing the questionnaire, the respondents placed their questionnaires in unmarked envelopes and handed them to the consultant, who forwarded the envelopes to a document management company that scanned the questionnaires and delivered the
output data to the evaluator.\textsuperscript{25} The questionnaire was designed so that it would meet the technical requirements for optical scanning and rules were created for the accurate coding of variables. Prior to the study, the questionnaire was tested on a sample of ten co-workers at SoRAD and was sent for scanning. The data set delivered on the basis of this sample was then double checked against the completed questionnaires to ensure that the coding had been carried out in accordance with the evaluator’s specifications.

Data collection was quite complicated given the large sample sizes, the fact that the study population was geographically distributed across the whole of the country and also the fact that the PfL program delivery was spread over a one-year period. Staff consultants at the regiments had to keep track of when the PfL program was scheduled and ensure the on-time administration of questionnaires to participants for the pre- and post-tests. To help facilitate data collection, the consultants continuously filled in evaluation reports in connection with “questionnaire sessions” and sent these to the project leader. These reports contained observations and questions raised during data collection as well as the names of participants, so that the follow-up questionnaires for the 12-month follow-up study could be administered to participants at the correct time.

5.3 Quality of outcome data

To ensure the quality and relevance of the questionnaires, a pilot study was conducted with a strategically selected sample of eight people from the Swedish Armed Forces. The respondents represented unions, a representative of the human resources unit at military headquarters, a legal expert at military headquarters, representatives for alcohol and drug prevention within the organization and also a representative of the administrative functions at military headquarters. Their feedback on the questionnaires helped to improve the layout and content of the survey, but the pilot study was also seen as an important means of obtaining approval for the evaluation within the organization.

\textsuperscript{25} The scanning of questionnaires was conducted by the document and records management company Recall in Östersund.
5.3.1 The experiment group

The experiment group population consisted of 848 individuals who participated in PfL training in the course of 2005. Only those individuals who participated in both the pre- and post-test were included in the evaluation. Each individual was identified and matched between T1 and T2 using a self-generated identification number that was cross-checked against both demographic information and the response date entered on the questionnaire by the respondent. 792 of the respondents had provided sufficient information to facilitate the matching procedure. In the follow-up study 12 months after the program, 520 individuals participated. Out of these, pre- and post-test answers were successfully matched for 446 individuals. In total, 446 out of 792 respondents (56%) of the experiment group qualified for participation in the evaluation. Due to the relatively large drop-out rate between T1 and T2, a simple analysis of the differences between participants in the evaluation and drop-outs has been conducted. In the following tables (5-6), those in the experiment group who are included in the study are compared to those individuals who participated in the study at T1 but not in the one-year follow-up study (evaluation drop-outs).

Table 8: Comparison of background data between the experiment group (N=446) and evaluation drop-outs (N=346). Swedish Armed Forces. Percent.

<table>
<thead>
<tr>
<th></th>
<th>Experiment group</th>
<th>Drop-outs</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>22</td>
<td>30</td>
<td>***</td>
</tr>
<tr>
<td>Men</td>
<td>78</td>
<td>70</td>
<td>***</td>
</tr>
<tr>
<td>Air Force</td>
<td>5</td>
<td>2</td>
<td>**</td>
</tr>
<tr>
<td>Army</td>
<td>53</td>
<td>59</td>
<td>-</td>
</tr>
<tr>
<td>Navy</td>
<td>26</td>
<td>10</td>
<td>***</td>
</tr>
<tr>
<td>Armed Forces Logistics/support functions</td>
<td>16</td>
<td>29</td>
<td>***</td>
</tr>
<tr>
<td>Born before 1966</td>
<td>47</td>
<td>47</td>
<td>-</td>
</tr>
<tr>
<td>Born 1966 or later</td>
<td>53</td>
<td>53</td>
<td>-</td>
</tr>
</tbody>
</table>

**p<0.05 ***p<0.01
Results indicate that there are statistically significant differences between participants in the evaluation and the drop-outs in relation to gender and type of employment and also the branch of the armed forces at which the participants were employed.

<table>
<thead>
<tr>
<th></th>
<th>Experiment group</th>
<th>Drop-outs</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk awareness</td>
<td>19.38</td>
<td>19.09</td>
<td>-</td>
</tr>
<tr>
<td>Drinks on a typical day drinking alcohol</td>
<td>2.04</td>
<td>2.06</td>
<td>-</td>
</tr>
<tr>
<td>How often alcohol consumption/month</td>
<td>1.87</td>
<td>1.76</td>
<td>**</td>
</tr>
<tr>
<td>Frequency binge drinking</td>
<td>0.97</td>
<td>0.97</td>
<td>-</td>
</tr>
<tr>
<td>Total AUDIT-score</td>
<td>5.19</td>
<td>5.08</td>
<td>-</td>
</tr>
</tbody>
</table>

**p<0.05

There was a statistically significant difference in how often alcohol was consumed between participants and drop-outs.

The conclusion about the quality of the data in the experiment group was that even though less than 60 percent of those Prime for Life participants who did the pre-test (T1) were included in the evaluation, there seemed to be no alarming differences between participants and drop-outs when it came to the outcome variables (i.e. risk awareness and alcohol behaviour) at T1. At the same time, the analysis revealed significant differences between evaluation participants and drop-outs when it came to gender distribution and which branch of the armed they represented.
5.3.2 The comparison group

In the comparison group, 1,857 individuals (of the target of 3,000) participated in the pre-test in 2005. In the 12 month follow-up, 515 individuals participated. 480 individuals where then successfully matched and included in the study. This means that only 26 percent of the total number of participants in the pre-test was included in the evaluation.

I have conducted an analysis of the differences between participants in the evaluation and drop-outs in the comparison group. As with the experiment group, the answers of those who had participated in the pre-test but were not included in the evaluation (N=1,359) were compared with those of the participants in the study (N=480).

<table>
<thead>
<tr>
<th></th>
<th>Comparison group</th>
<th>Drop-outs</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>18</td>
<td>12</td>
<td>***</td>
</tr>
<tr>
<td>Men</td>
<td>82</td>
<td>87</td>
<td>***</td>
</tr>
<tr>
<td>Air Force</td>
<td>23</td>
<td>43</td>
<td>***</td>
</tr>
<tr>
<td>Army</td>
<td>70</td>
<td>37</td>
<td>***</td>
</tr>
<tr>
<td>Navy</td>
<td>0</td>
<td>20</td>
<td>***</td>
</tr>
<tr>
<td>Armed Forces Logistics/support functions</td>
<td>6</td>
<td>1</td>
<td>***</td>
</tr>
<tr>
<td>Born before 1966</td>
<td>56</td>
<td>38</td>
<td>***</td>
</tr>
<tr>
<td>Born 1966 or later</td>
<td>44</td>
<td>62</td>
<td>***</td>
</tr>
</tbody>
</table>

***p<0.01

Background data on comparison group participants and drop-outs at T2 reveal significant differences between evaluation participants and drop-outs.
### Table 11: Comparison of Outcome Variables. Mean Scores for Comparison Group (N=480) and Evaluation Drop-outs (N=1359) at T1. Comparison Group, Swedish Armed Forces.

<table>
<thead>
<tr>
<th></th>
<th>Comparison group</th>
<th>Drop-outs</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk awareness</td>
<td>18.49</td>
<td>18.71</td>
<td>-</td>
</tr>
<tr>
<td>Drinks on a typical day</td>
<td>1.90</td>
<td>2.03</td>
<td>**</td>
</tr>
<tr>
<td>drinking alcohol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often alcohol consumption</td>
<td>1.94</td>
<td>1.87</td>
<td>**</td>
</tr>
<tr>
<td>/month</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency binge drinking</td>
<td>0.91</td>
<td>1.02</td>
<td>***</td>
</tr>
<tr>
<td>Total AUDIT-score</td>
<td>4.66</td>
<td>5.05</td>
<td>**</td>
</tr>
</tbody>
</table>

**p<0.05 ***p<0.01

As Table 11 (above) shows there was a statistically significant difference between the evaluation participants and drop-outs in the comparison group on all alcohol-related outcome variables. The drop-outs had significantly higher scores on all variables except on how often alcohol was consumed.

The conclusion of this simple analysis is that the drop-outs in the comparison group differed from the evaluation participants and that this has affected the quality of the data in the study. The evaluation participants in the comparison group are not representative of the population that was sampled for the study.

5.3.3 Comparing the experiment and comparison groups at T1

In the following, the evaluation experiment and comparison groups are compared at T1 to analyse whether there were any statistically significant differences between the two.
Table 12: Comparison between experiment (N=446) and comparison group (N=480). Swedish Armed Forces. Percent

<table>
<thead>
<tr>
<th>Branch</th>
<th>Experiment group</th>
<th>Comparison group</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>22</td>
<td>18</td>
<td>-</td>
</tr>
<tr>
<td>Men</td>
<td>78</td>
<td>83</td>
<td>-</td>
</tr>
<tr>
<td>Air Force</td>
<td>5</td>
<td>23</td>
<td>***</td>
</tr>
<tr>
<td>Army</td>
<td>53</td>
<td>70</td>
<td>***</td>
</tr>
<tr>
<td>Navy</td>
<td>25</td>
<td>0</td>
<td>***</td>
</tr>
<tr>
<td>Armed Force Logistics/support</td>
<td>17</td>
<td>6</td>
<td>***</td>
</tr>
<tr>
<td>Born before 1966</td>
<td>77</td>
<td>65</td>
<td>***</td>
</tr>
<tr>
<td>Born 1966 or later</td>
<td>23</td>
<td>35</td>
<td>***</td>
</tr>
</tbody>
</table>

***p<0.01

As Table 12 (above) indicates, there were statistically significant differences in the distribution of participants between the experiment and comparison groups at T1 with regard to the branch of the armed forces they represented and age.
TABLE 13: COMPARISON BETWEEN EXPERIMENT (N=446) AND COMPARISON GROUP (N=480) AT T1, SWEDISH ARMED FORCES. MEAN SCORES.

<table>
<thead>
<tr>
<th></th>
<th>Experiment group</th>
<th>Comparison group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk awareness</td>
<td>19.36</td>
<td>18.48</td>
<td>**</td>
</tr>
<tr>
<td>Drinks on a typical day drinking alcohol</td>
<td>2.04</td>
<td>1.89</td>
<td>**</td>
</tr>
<tr>
<td>How often alcohol consumption/month</td>
<td>1.87</td>
<td>1.94</td>
<td>-</td>
</tr>
<tr>
<td>Frequency binge drinking</td>
<td>0.98</td>
<td>0.90</td>
<td>-</td>
</tr>
<tr>
<td>Total AUDIT-score</td>
<td>5.18</td>
<td>4.63</td>
<td>**</td>
</tr>
</tbody>
</table>

**p<0.05

The mean values of the key outcome valuables in the evaluation were higher in the experiment group in all cases except for the variable measuring how often alcohol was consumed. At T1, there was statistically significant higher risk awareness in the experiment group, and significant differences in how many drinks were consumed on a typical drinking day and on the mean AUDIT-score.

5.3.4 Conclusions about data quality

Both the experiment and the comparison groups in the evaluation of Prime for Life within the Swedish Armed Forces had large drop-out rates between the pre- and post-tests. This makes the results of the evaluation of PfL within the Swedish Armed Forces less reliable than would be desired.
5.4 Short-term program evaluation results

Of the 849 individuals who participated in the PfL program, 686 (80 %) completed the short-term course-evaluation questionnaire distributed after the program.

<table>
<thead>
<tr>
<th></th>
<th>Overall impression of the program</th>
<th>The instructors’ teaching style</th>
<th>The instructors’ knowledge</th>
<th>The program textbook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
<td>33</td>
<td>42</td>
<td>54</td>
<td>17</td>
</tr>
<tr>
<td>Good</td>
<td>54</td>
<td>50</td>
<td>42</td>
<td>66</td>
</tr>
<tr>
<td>Bad</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Very bad</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Don’t know/Non-response</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

The Swedish Armed Forces personnel participating in the PfL program gave it very positive ratings, from 83 percent rating the program as good or very good when it came to the program textbook to 96 percent rating the instructors’ knowledge in the field as good or very good.

5.5 Main outcome results of the evaluation

In the following presentation of the outcome results from the evaluation of Prime for Life in the Swedish Armed Forces, results have been adjusted for pre-test differences for each outcome variable, age and military branch.
Table 15: Effects of Prime for Life (PFL) on key outcomes 12 months after the program (T2). Results from analyses of covariance (ANCOVA), adjusting for pre-test differences in outcome, age and military branch (full sample, 95% confidence intervals in parentheses). Swedish Armed Forces.

<table>
<thead>
<tr>
<th>Variable \ outcome</th>
<th>Risk awareness</th>
<th>No. of drinks on a typical day drinking</th>
<th>How often alcohol consumption</th>
<th>Frequency binge drinking</th>
<th>Total AUDIT score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
</tr>
<tr>
<td>Pfl</td>
<td>5.31</td>
<td>1.10</td>
<td>0.11</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>[4.63 – 5.98]</td>
<td>[0.96 – 1.24]</td>
<td>[0.95 – 0.17]</td>
<td>[-0.02 – 0.24]</td>
<td>[-0.04 – 0.16]</td>
</tr>
</tbody>
</table>

| No. of observations | n<sub>total</sub>= 911 | n<sub>PFL</sub>= 439 | n<sub>control</sub>= 472 | n<sub>total</sub>= 883 | n<sub>PFL</sub>= 426 | n<sub>control</sub>= 457 | n<sub>total</sub>= 891 | n<sub>PFL</sub>= 429 | n<sub>control</sub>= 462 | n<sub>total</sub>= 879 | n<sub>PFL</sub>= 422 | n<sub>control</sub>= 457 | n<sub>total</sub>= 865 | n<sub>PFL</sub>= 415 | n<sub>control</sub>= 450 |

SMD=Standardized Mean Difference (Cohen’s d).

As Table 15 illustrates, the effect size of the PFL program is large in relation to risk awareness. Outcome measures on alcohol behaviour indicate trivial increases on all alcohol behaviour outcome variables. Results are similar when participants are divided into men and women, i.e. there is a large effect on risk awareness but trivial effects on alcohol behaviour (see Tables 3.1 and 3.2, Appendix 3).
5.5.1 Outcomes for risk consumers

In the following, women with an AUDIT-score of 6 or higher and men with an AUDIT-score of 8 or higher have been selected for inclusion in the analyses.\(^{26}\)

**Table 16: Effects of Prime For Life (PFL) on key outcomes 12 months after the program (T2). Results from analyses of covariance (ANCOVA), adjusting for pre-test differences in outcome, age and military branch (full sample, 95% confidence intervals in parentheses). Risk consumers, Swedish Armed Forces.**

<table>
<thead>
<tr>
<th>Variable \ outcome</th>
<th>Risk awareness</th>
<th>No. of drinks on a typical day drinking</th>
<th>How often alcohol consumption</th>
<th>Frequency binge drinking</th>
<th>Total AUDIT score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
</tr>
<tr>
<td>PFL</td>
<td>5.66</td>
<td>1.51</td>
<td>-0.43</td>
<td>-0.35</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>[4.35 – 6.96]</td>
<td>[1.20 – 1.82]</td>
<td>[-0.73 – -0.10]</td>
<td>[-0.66 – -0.04]</td>
<td>[-0.14 – 0.24]</td>
</tr>
</tbody>
</table>

| No. of observations | n\(_{total}\)= 168 | n\(_{PFL}\)= 101 | n\(_{control}\)= 67 | n\(_{total}\)= 166 | n\(_{PFL}\)= 101 | n\(_{control}\)= 65 | n\(_{total}\)= 166 | n\(_{PFL}\)= 101 | n\(_{control}\)= 65 | n\(_{total}\)= 168 | n\(_{PFL}\)= 101 | n\(_{control}\)= 67 |

SMD=STANDARDIZED MEAN DIFFERENCE (COHEN’S D).

As Table 16 (above) illustrates, there was a large impact of the program on risk awareness among risk consumers, whereas the effect in relation to alcohol behaviour varies between trivial and small.

---

\(^{26}\) Due to the low number of women in the analysis (8 in the experiment group and 9 in the comparison group), results are not presented by gender (men and women).
5.5.2 Comments regarding the outcome results

The outcome evaluation of Prime for Life was subject to large drop-out rates which have most likely caused a selection bias. Even though results have been adjusted for pre-test differences for each outcome variable, age and military branch, they should be interpreted with caution.

The results are similar to those of the evaluation of Prime for Life at Örebro University, i.e. the program’s effect on risk awareness is large, whereas the effects on alcohol behaviour are trivial or small.
6. Generated knowledge and possible policy implications

Knowledge from an evaluation may be either instrumental, when a decision or action follows from an evaluation, or conceptual, when the use of evaluation influences thinking and expands understanding by increasing knowledge (Patton 2001). It could be argued that the EBP movement adheres primarily to a more instrumental view of knowledge, since politicians and professionals are supposed to make rational decisions (and achieve accountability) through the application of “sound evidence” (Biesta 2007, Sanderson 2002a, Sanderson 2002b).

Evidence-based policy tends to focus on accountability in terms of the results of social activities – there is a need to know that government is working effectively (Sanderson 2002b). As discussed in Chapter 2, evidence-based policy making views decision making as a rational process with the goal of seeking a way, and ideally the best way, to improve society. If policy is goal-oriented, then it also makes sense that evaluations are goal-oriented.

My own view of knowledge of the social world is that it is socially constructed and that the role of knowledge in policy making is far more complex than this. Research does not instrumentally inform policy development; rather it reaches decision-makers in unsystematic and diffuse forms (Sanderson 2002b). Unintended and unanticipated consequences of social activities undermine our ability to predict and control the basis of knowledge. My point here is not to say that systematic knowledge about social activities does not exist, rather that there needs to be a balance between over-optimistic “objectivist claims” and the over-pessimistic nihilism of relativists (Trigg 2001).

To expand on the lessons learned from my evaluation, I will in the following try to go beyond the outcome measurements that to a large extent shaped my evaluation studies and the reporting of their results. I will exemplify different kinds of knowledge generated by the evaluations and discuss what kind of policy implications they might lead to.

I will start with the main outcome results of the evaluations – what kind of information do these results provide policy makers with about the Prime for Life program, and what sort of policy-decisions do they suggest? I then will try to go beyond the “what works” perspective and discuss what other knowledge the evaluations have generated and in which ways this
information can contribute to the development of social policy, and improve evaluations of such activities.

6.1 Outcome results that (for the most part) confirm previous research results

The results from both my studies confirm the results of previous Swedish evaluation studies showing that the PfL program has a large impact on risk awareness but that it is uncertain whether it has the desired effects on alcohol use (Hallgren et al. 2009, Leifman 2008, Sjölund & Andréasson 2004).

The overall conclusion from the two evaluations is that the general effect of the PfL program on risk awareness is large, but that the primary alcohol prevention effect is, at best, limited. The evaluations found no clear indications of positive changes in alcohol behaviour among participants in general. These results are very much in line with the substantial body of international research which indicates that education-based alcohol and drug prevention programs have either indeterminate or limited effects on the behaviour it is intended to change (Babor et al. 2010, Baker et al. 2006, Cuijpers 2002, Foxcroft et al. 2003, WHO 2004, WHO 2009).

For risk consumers of alcohol, the results also indicate a large effect of the program on risk awareness, but only small effects on alcohol behaviours. The study at Örebro University showed a large effect on the mean AUDIT-score, and indicated that the large effect on the mean AUDIT-score for risk consumers was explained by a decrease between T1 and T2 in experienced alcohol-related harms such as black-outs or alcohol-related injuries. Though these results are in line with the theoretical assumptions of Prime for Life, they are not validated by the results from other evaluations of PfL in Sweden (i.e. Hallgren et al. 2009, Leifman 2008, Sjölund & Andréasson 2004). In the evaluation of PfL among students at high-schools in Stockholm (N=602), the evaluators noted that five months after the program there was a “(...) a major decrease in consumption between the pre-test and the post-test” for high-risk consumers (…)” (Sjölund & Andréasson 2004, 31. Author’s translation). Because of the small group of risk consumers (n=78) the results were not statistically significant and the result was disregarded. The evaluators deliberately chose not to report the effect size for this change (Sjölund & Andréasson 2004). In the 20 month follow-up, the outcome for risk
consumers was reported on the basis of a combined mean score on the first three questions of the AUDIT questionnaire. Results then indicated no difference in alcohol behaviour between risk consumers and the group of Pfl participants as a whole (Leifman 2008).

6.1.1. Policy implications of the outcome results

The outcome of the Prime for Life program on risk awareness is evident and its content and delivery was highly appreciated in the course evaluations. The program can be seen as an effective group-based intervention for increasing awareness of the risks associated with alcohol.

However, from an alcohol policy perspective, the recommendation from the evaluations for decision-makers would be not to use the Pfl program when looking for a primary prevention program that is intended to have a direct impact on participants’ alcohol behaviour.

Personally, I would be cautious about ruling out the possibility that the program might actually have an impact on alcohol behaviour among risk consumers of alcohol. The reason for this is not so much the large effect on the mean AUDIT-score identified in the Örebro University study; but I do think the program needs to be implemented and studied in Sweden as an intervention aimed specifically at risk consumers – rather than as a primary prevention effort in which risk consumers are identified and analysed ad hoc. (I will develop this argument in Chapter 6.2.2 below.)

---

27 I.e. how often alcohol was consumed, how many units of alcohol were consumed in one setting and the frequency of binge drinking.
6.2 Increased understanding of the value of quality in the implementation of social interventions.

Implementation analyses can complement outcome studies by providing information that may help to improve policy and practice. In the following, implementation refers to a set of qualities (such as planning, obtaining approval for, promoting and executing an activity in an organizational context) which go beyond whether or not the specific intervention in itself is well documented (see Chapter 8.1.3).

The primary focus in the two PfL evaluations presented in this thesis has been on outcomes and on assessing the effectiveness of the PfL program in relation to alcohol consumption. However, the studies clearly illustrate that implementation is important for understanding what prospects it produces for long-term alcohol prevention:

The evaluation at Örebro University showed that the staff at the department felt that the decision to implement the program had come as a top-down decision. This was a problem for them as they felt that they had little opportunity to influence the nature and content of the intervention. However, it does not seem that the staff’s generally negative attitude towards the program affected how those students who chose to participate in the program experienced it (they were very satisfied with it). However, it is possible that the level of participation in the program would have been higher if the staff had promoted it in a more active and positive way.

In addition, the qualitative data from the evaluation showed the complexity of having alcohol as a teaching subject in an academic environment that itself has strong traditions associated with alcohol consumption (prop. 2005/06:30). While staff at the department talked about their student environment as unique and spoke of “their” students as having learned a more sophisticated way of consuming alcohol (compared to students in other parts of the university), interviews with students and the evaluation outcome data showed that the reality was more complex, with high levels of heavy drinking among students at the
department in parallel with a “proclaimed” culture of drinking alcohol with an eye to “quality rather than quantity.”

Among departmental staff there was a certain level of scepticism towards the program and there were also contradictions in the way the relationship to alcohol at the department was described, which led me to feel that the resistance towards the alcohol prevention initiative may partly have been due to the staff feeling that the department had been targeted and labelled as a problematic drinking environment by the Alcohol Committee and the Örebro University management. If this was the case, it strengthens the case for the view that the planning and execution of the Prime for Life program was not ideally managed.

The evaluation of Prime for Life in the Swedish Armed Forces did not include any systematic collection of data focusing on the implementation of the program and the results previously presented (in Chapter 5) focused solely on the program’s outcome results. The Swedish Armed Forces specifically wanted to limit the study to look at outcomes on alcohol consumption among participants, even though the implementation of the program was not conducted in the absence of known problems:

In 2004, the Swedish parliament approved a reorganization of the Swedish Armed Forces that included the disbanding of a number of regiments and a reduction in staff numbers (both military and civilian personnel) of approximately 25 percent (from about 20,000 employees). As the evaluation results revealed, the Swedish Armed Forces did not manage to achieve its goals in relation to the implementation of Prime for Life during the evaluation period, and there was a substantial drop-out rate in the evaluation. When I contacted one particular regimental commander to follow up on the implementation of PfL, since it seemed to have come to a halt at his regiment, he simply said: “I have other things to prioritize... Like giving people notice.” Written field reports from staff consultants who administered the questionnaires at different regiments also testified to the fact that some regiments were unwilling to put time into planning, implementing and evaluating a program that had been “forced upon” them while they were at the same time having to deal with the major changes that were occurring within the organization.

It could be argued that, at the time of the evaluation, the need for an alcohol prevention program within the Swedish Armed Forces was perhaps greater than ever: Stress due to
organizational changes could be a significant contributor to alcohol consumption (Pohorecky 1991, Vasse, Nijhuis & Kok 1998). It might be reasonable for an organisation’s management to implement programs and other activities that are considered likely to prevent problems that may result from a re-organization, for example. It may therefore have been a deliberate choice of the Supreme Commander to maintain programs and policies related to the health and wellbeing of the personnel at this time. However, the support for this decision in the organization as a whole was probably not as strong as would be desirable. All in all, I think it is reasonable to conclude that the situation within the Swedish Armed Forces and the resistance at certain regiments to participating in and evaluating the PfL program complicated both the implementation of the program and the evaluation study.

6.2.1 Policy implications of observations regarding implementation quality

Issues regarding implementation, including different stakeholder perspectives, are often not addressed in outcome evaluations (Chen & Garbe 2011). Given that stakeholders such as the teachers and administrative staff at Örebro University or the regimental commanders within the Swedish Armed Forces are the people who are responsible for delivering services to their students and staff on a day-to-day basis, I think that their concerns about the attractiveness of an intervention and their capacity to manage its practical implementation are of relevance. If the EBP movement does not adequately address practical issues, stakeholders might not find evidence-based interventions useful (Wanderman et al. 2008).

The information that is available on the implementation of PfL at Örebro University and in the Swedish Armed Forces points to the importance of gaining acceptance for an intervention by listening to different stakeholders and of planning the intervention in a manner that considers their interests and opportunities to both support and become engaged in the intervention. Failure to properly implement an activity may have a negative impact on the possibility of reaching desired outcomes, and poor project management may lead to a significant waste of money and effort and may cause client dissatisfaction that damages long-term opportunities for alcohol prevention policies. Top-down and expert-initiated randomized controlled studies have a shortcoming here in the sense that the design often limits the opportunities for participants and stakeholders to themselves choose what is
relevant for them and to have a say in how and when an intervention is chosen and implemented. From this perspective, I think it is important for decision-makers to consider models for implementing and evaluating efforts that acknowledge different stakeholder values and include their perspective in the overall assessment of the worth and merit of an intervention.

6.2.2 A note on the limited knowledge on the risks of mixing high- and low-risk participants

Another aspect of implementation is the matching of interventions to the personal characteristics of participants. Such matching is seldom carried out. Even in the treatment of offenders, it is quite common that treatment programs tend to be implemented with a “one size fits all” approach (Matthews et al. 2001, 466).

University students are commonly described as a “high-risk group” when it comes to alcohol consumption (Bullock 2004, Crawford & Novak 2010, prop. 2005/06:30). In the evaluation at Örebro University, about 40 percent of the participating students had a risky alcohol consumption as defined by the AUDIT-10 scale compared to approximately 18 percent in the Swedish Armed Forces (Sandberg 2007b) and about 8 percent in the general population in 2005 (Boman et al. 2006). Even in a high-risk setting, such as a university, it is important to note that the majority of students are not high-risk consumers of alcohol. Since the PfL program was originally designed for groups that make high-risk choices, it is important to be aware that when implemented as primary prevention, it is not reasonable to expect all participants (at least not in the short-term) to change their alcohol behaviour as a direct consequence of program participation. On the contrary, many participants might actually perceive their alcohol consumption as low-risk, and see no immediate reason to change their alcohol behaviour at all. In an outcome evaluation of alcohol prevention, low-risk consumers of alcohol might actually disguise the impact of the program when it comes to alcohol consumption:
Many programs appear ineffective when it may be that inappropriate cases are masking the success that these programs are having with the clients who are appropriate for the services provided (Matthews et al. 2001, 466-467).

This problem can quite easily be solved by studying high- and low-risk participants separately in an evaluation, as was done in my evaluations of Prime for Life. However, the likelihood of finding statistically significant results is dependent on the size of the population studied, and if, for example, the potential success of a program is restricted to a minority of participants, outcomes may not yield significant results unless the population is very large. Some outcome evaluators do not report on effect sizes for outcome variables unless changes are statistically significant (see for example Sjölund & Andréasson 2004).

Controlling for high- and low-risk groups when analysing evaluation data can illuminate differences between groups, but it does not compensate for the possible impact that the actual mixing of high- and low-risk participants in the program may have on its implementation and outcome. To my knowledge, no studies have been conducted of the impact that the lack of matching interventions with client characteristics has on alcohol prevention initiatives. However, in criminal justice programs, it has been shown that a failure to match the level of service to the risk-level of the offenders can have serious consequences – the provision of intensive services and controls to low-risk offenders may be harmful, since it interferes with the generally pro-social lifestyles of these offenders (Matthews et al. 2001).

6.3 A better understanding of an intervention and its outcomes – the value of a program-theoretical perspective

It is not required for an outcome evaluator to understand how an intervention program works in order to estimate its net effect. “A focus on what works alone produces a ‘one dimensional’ approach to measuring outcomes that concentrates on one variable at the expense of many other independent variables” (Cook 2006, 10). Alcohol behaviour and
related problems are dependent on numerous factors such as psychological and genetic factors, environmental variables and lifestyle. It is not likely that a given intervention can successfully target all the different elements that help to explain the different levels of alcohol consumption among the members of a given group. The use of a program theory approach can complement outcome studies by providing hypotheses about how different components in the implementation of a program can contribute to different outcomes.

In my evaluations, I used a simple program theoretical approach in order to better understand the study object, to structure analysis and to help me balance the results by not drawing overly far-reaching conclusions about the program on the basis of the outcomes. The program theoretical approach also allowed for a deeper analysis of the program and the evaluation results.
<table>
<thead>
<tr>
<th>Program theory step</th>
<th>Swedish Armed Forces</th>
<th>Örebro University</th>
</tr>
</thead>
</table>
| 1. Implementation   | Overall program planning and approval:  
Non-satisfactory; implementation did not achieve goals for program and evaluation  
PfL content and delivery:  
High level of satisfaction among program participants | Overall program planning and approval:  
Non-satisfactory; scepticism and resistance towards the program among departmental staff.  
PfL content and delivery:  
High level of satisfaction among participating students |
| 2. Participants acquire knowledge and accept it | High level of trust for instructors and the program content among participants | High level of trust for instructors and the program content among participants |
| 3. Increased risk awareness | Large ES on risk awareness 12 months after program participation | Large ES on risk awareness 12 months after program participation |
| 4a) Changed alcohol behaviour for risky alcohol consumers | Results indicating mostly trivial effects | Trivial to moderate effects |
| 4b) Changed alcohol behaviour for all participants | Inconclusive results indicating mostly trivial effects | No effect to small effects of the program |

**Figure 3: Examples of results regarding the Prime for Life program**

Figure 3 above provides an overview of observations regarding the different steps in the program theory (see Chapter 3). In the following, I will go through my observations regarding each step more in detail.
6.3.1 Implementation

In short, the results from the evaluations indicated that implementation went well from the users’ (i.e. the program participants’) perspective; both the students and military personnel who had participated in the PfL program were very satisfied with the program content and delivery.

However, as was illustrated above (Chapter 6.2), implementation was problematic in the Swedish Armed Forces, since the organisation was unable to achieve its goals regarding the implementation of the program or the collection of data for the evaluation. It is clear that within the Swedish Armed Forces, the problem was not so much about the PfL program per se not being accepted by stakeholders within the organisation, as it was about the organisation’s inability to implement the program as planned during the evaluation period. It seems very likely that the timing of the study was less than perfect. At Örebro University, there was scepticism about the program among departmental teachers and administrative staff. This was most likely caused by a lack of communication between the university management and the departmental staff, who felt that the program had been chosen and implemented without their approval or involvement.

While the implementational problems identified in the studies may not have affected the outcome of the PfL program, it is quite obvious that a greater level of support for the PfL program among staff, both at Örebro University and in the Swedish Armed Forces, would have been necessary for the program to become a credible part of a systematic work-environment effort (Arbetsmiljöverket 2003). At Örebro University in particular, where I have systematically collected data that indicate a solid resistance towards the program among departmental staff, it seems unlikely that the PfL program could be seen as a long-term alcohol prevention effort.

---

28 As illustrated in Chapter 5, my interviews clearly indicated that the staff at the Department of Hospitality, Culinary Arts and Meal Science were not satisfied with the content and delivery of the actual PfL program. It should be noted that this observation is based on my sample of interviewed representatives of the department and the evaluation did not involve studying their actual participation in the program and what possible impact the program may have had on their risk awareness or alcohol behaviours.
6.3.2 Received knowledge and risk awareness

When it comes to alcohol problems, many prevention programs are based on educational services. The fundamental assumption (which is rarely made explicit) is that people will change their behaviour if they are given information about the reasons for doing so (Rossi, Lipsey & Freeman 2004). However, the program theory of PfL stipulates that for the program to work, participants need to find the program content trustworthy and useful. Available data indicate that participants in the PfL program had high levels of trust in the instructors, the course content and the information provided during the program. The most apparent result of both evaluations has been a large ES of the PfL program on risk awareness. The risk perspective taught during the PfL program has been learned and remembered by program participants.

6.3.3 Changes in alcohol behaviour

The program theory of the PfL program can help to explain why primary prevention outcomes could not be found in the evaluations of the program, since the theoretical foundation of Prime for Life makes it quite clear that the program was not originally designed to change alcohol behaviour among participants with no signs of risky alcohol consumption.

In the Örebro study, there is a slight tendency towards more positive effects of the program among risk consumers, with a moderate decrease being noted in the mean AUDIT-score for risk consumers. However, the results from the study in the Swedish Armed Forces provide no evidence to suggest a positive change in alcohol behaviour among risk consumers.

Given that we do not know whether the mixing of high- and low-risk participants might affect the impact of the program (see Chapter 6.2.2 above), I cannot draw any clear conclusions about whether or not the program design is unlikely to produce desired outcomes – a theory error – or whether the activity is not being executed in a way that initiates the mechanisms that are expected to lead to desired outcomes, so called implementation error (Sandberg & Faugert 2012, Statskontoret 2001). To shed light on this
question, further studies of the program are needed, particularly studies of the program being implemented as a secondary and tertiary prevention initiative, which is what it was originally designed as.

### 6.3.4 Policy implications from the program-theoretical perspective

Until further studies have been conducted of Prime for Life as a program designed for those who make high-risk choices, it is important to bear in mind that in my own studies, the program has been both implemented and evaluated as a primary prevention initiative.

In my opinion, the way that the PfL program has been put into action reveals a discrepancy between the program itself and the fact that it is being promoted, used and evaluated as primary alcohol prevention. There is not necessarily anything wrong in this, as long as the potential value and merit of the program is being understood properly by stakeholders. It is just as important that the program itself is assessed in a balanced way by evaluators. If the program design does not make it reasonable to expect any major effects on alcohol behaviour among participants when the program is implemented as primary prevention, I think it can be questioned what value there is in actually assessing the program from this perspective.

In order to learn whether the failure of an intervention to produce desired results is due to flaws in the program or because it was poorly implemented is practically impossible in an evaluation that focuses solely on intervention outcomes. For this reason I think it is necessary in evaluations in the field of alcohol prevention to make some kind of effort to try to distinguish between a failure of the program design and a failure in the implementation of the program in question. Meta-analyses of interventions provide an opportunity to produce statistically reliable information about a certain type of action. If evidence then points to an intervention not being effective, the conclusion is generally that it “does not work”. Studies of correctional programs have shown however that even programs based on strong theoretical models (such as PfL) or even on evidence-based practices are often inconsistently implemented and lack structure (Latessa 1999). When it comes to improving alcohol policy, I do not think that a steady supply of evidence-based interventions is enough to improve
practice. A broader, systematic program approach to evaluations is needed so that analysis can move beyond simplistic linear reasoning (i.e. “intervention X causes result Z”) and can genuinely recognize what is required to implement activities in an effective way.
7. Shortcomings and strengths of my evaluation studies

In this chapter, I will discuss some of both the shortcomings and the strengths of the evaluation studies that I have reported in this thesis. I will start by assessing the quality of my outcome studies by comparing them to the “golden standard” for outcome studies in EBP. I will then discuss in more detail the narrow outcome-focus of outcome evaluations and the top-down perspective that they usually imply. Finally, I will briefly discuss what I could have done, and would do, differently if I had the opportunity to conduct the studies again.

7.1 Comparing my evaluations to “the golden standard”

In this section I will review the quality of my studies using the “ideal” from an evidence-based policy perspective as a benchmark. In the sections on “study design”, “specificity of the data”, “replicable intervention” and “outcome measures that target what the intervention aims to change”, I will describe the National Board of Health and Welfare’s criteria for quality in outcome studies (derived from Socialstyrelsen 2009b, 2010b) and then assess the quality of my own evaluations based on these criteria.

7.1.1 Study design

The two kinds of outcome evaluation design that provide the most reliable conclusions about the effects of interventions are comparative studies with pre-and post-tests and randomized controlled studies (RCT).

My studies were comparative outcome evaluations with pre- and post-tests. However, they were not randomized, which leads to less statistical certainty and lowers the value of an evaluation for evidence-based practice. As was discussed in Chapter 2, randomization can be very difficult to accomplish when evaluating social interventions, and it is sometimes associated with problems that involve a direct conflict with the interests of different stakeholders who are involved in the activity being studied. For example, the Alcohol
Committee and Örebro University specifically wanted to offer the PfL program to all students at the Department of Hospitality, Culinary Arts and Meal Science, and they had already started the project when the evaluation was initiated. It is my belief that insisting on using a randomized study design would have compromised stakeholder interests and (even further) undermined the implementation process. I therefore chose a practical solution that meant minimizing my own control over the experiment rather than restricting the activity of those involved in it (for example the Alcohol Committee and staff and students at the department). When it comes to the evaluation of PfL in the Swedish Armed Forces, I also made the decision to conduct a quasi-experimental study, although with the benefit of hindsight I think that a smaller, randomized study would have been preferable. I will develop this idea further below (“What could I have done differently?”; see Chapter 7.4).

7.1.2 Specificity of the data

For a study to be trustworthy, it is necessary that the information available for analysis has been collected using reliable methods. High specificity is often ascribed to standardised tests, medical observations and register studies.

My studies used the AUDIT-scale to assess risky alcohol consumption, while alcohol-related variables that have been used in multiple studies at the Centre for Social Research on Alcohol and Drugs and in other research studies were used to address alcohol consumption and alcohol-associated problems. As was described in Chapter 4, methodological issues and questions used in the data collection process were also discussed with senior researchers at the Centre for Social Research on Alcohol and Drugs, and questionnaire content was tested on both fellow researchers and representatives of the survey target population prior to data collection.

7.1.3 Replicable intervention

According to the criteria for the quality of outcome studies, the studied intervention needs to be well documented in order to provide information that is as useful as possible to professionals. In this context, a well-documented intervention is one where the program
content can be identified in the evaluation and/or where references are made to a published manual.

The PfL program is a protocol- and manual-based intervention and I have both studied and referenced the PfL curriculum (PRI 2003) in my evaluations.

7.1.4 Outcome measures that target what the intervention aims to change

In meta-analysis, measures are often divided into “primary” and “secondary” outcome measurements. Primary measurements are those that measure what the treatment is meant to affect, while secondary measurements comprise all the other measurements used in an evaluation. According to the National Board of Health and Welfare (Socialstyrelsen 2009b, 2010b) a high-quality outcome study should use outcome measurements that target what the intervention aims to change (primary measurement).

I have applied what would undoubtedly be regarded as primary measurements in order to study the PfL program’s impact on risk awareness and alcohol behaviour. The outcome measures are variables that are also discussed in the PfL program, and they are commonly included in studies in the field of social research on alcohol and drugs (see for example Bullock 2004, Hallgren et al. 2009, Hradilova Selin 2003, Sjölund & Andréasson 2004).

I would like to problematize the notion that a given outcome measurement is labelled “primary” as it “measures what the program is meant to affect”. As discussed in Chapter 2, outcome evaluations can either pay attention to the intervention’s own goals or – as is common in outcome studies of alcohol prevention – what the evaluator defines as the intervention’s desired outcomes. Social researchers prefer to look for changes in alcohol consumption in the context of interventions and base their assessment of actions based upon this. At the same time, it is possible that what the researcher might consider to be a secondary measurement is actually the more desired outcome for another stakeholder. “(...) From a researcher’s standpoint these practical issues may be regarded as trivial, to stakeholders they are crucial” (Chen & Garbe 2011, 96). This might be a problematic gap between intervention research and “real-world practice” that the EBP movement needs to
better address in order to improve the transfer of knowledge to policy makers and professionals.

7.1.5 Well-conducted outcome evaluations – do my studies fit in?

From a strict methodological and design perspective, my studies do not possess all the qualities that are desired for an outcome evaluation to serve as a direct input to evidence-based practice (Socialstyrelsen 2010b). Nonetheless, they could most probably be included in a systematic review, although with a lower ranking than the “golden standard.” A meta-analysis that included my studies would most probably only use my outcome variables and would pay limited (or no) attention to the secondary measurements regarding, for example, the implementation of Prime for Life that were discussed in the previous chapter.

As has been shown in the presentation of results, the quality of the data collected in the evaluations was not as high as desired, and the problems of selection bias and large dropout rates are present in both studies. In this thesis, I have adjusted the results for pre-test differences in outcome and relevant background variables between experiment and control groups, but it is impossible to control for all the differences between the groups if non-random factors, such as motivation to participate, have disturbed the statistical analysis. The conclusions are in this sense a bit “shaky”, but they are validated by the findings from previous studies of the program in Sweden.

I have been hesitant to conduct statistical analyses to which I do not think the data are suited, particularly in the case of the Swedish Armed Forces study. I have also wanted to keep things as simple as possible, if for no other reason than to be able to present the results in a fashion that allows them to be easily read and understood by interested parties outside academia. However, I have striven to develop a deeper understanding of the Prime for Life program and of how it is supposed to work. This has helped me to better understand what I am evaluating and to look beyond the quantitative primary outcome measurements. I think this has given me an opportunity to structure both my analysis and the reporting of my results in what I believe is a fair and balanced way.
What strikes me, particularly when I compare my analysis with the other Swedish evaluations of Prime for Life, is that the program has repeatedly been implemented and assessed by scientific evaluators as a primary prevention program. This has been done in the absence of any theoretical or practical discussions about the possible implications of such an approach (see Chapter 6.2.2 above) and thus researchers have both planned and implemented the program and their evaluations in a way that is not necessarily in line with what the program was originally designed for – and that is not necessarily in tune with what stakeholders might expect from it.

The narrow focus on “primary measurements”, which I have also concentrated on in my own studies, in a sense adds more evidence to the stock of already well-established evidence showing that education-based alcohol and drug prevention programs have indeterminate or limited effects. The conclusions drawn from a strict outcome perspective do not help to explain why a prevention program such as PfL is still in use as a primary prevention effort despite all the evidence speaking against it. At this point, from a scientific perspective, it might be more interesting to find out why such programs are still in use.

For decision-makers, it might be more rational to implement and assess the PfL program and similar educational alcohol prevention as programs that build knowledge and that can be used for “policy formation”. I will discuss this further in the next section.

7.2 The narrow outcome focus

In outcome evaluations, the focus is often on a single outcome (Soydan & Vinnerljung 2002), and as has already been mentioned, when alcohol researchers evaluate primary alcohol prevention initiatives, they typically look at a reduction in alcohol consumption or alcohol-associated problems as the desired intervention outcome. Outcomes other than those pre-defined as the most relevant by the researcher are often ignored by outcome evaluators when summarising their results. Most crucial obviously, is when the narrow question “does the program work” fails to recognize that the program might have other outcomes that potentially cause more harm than good (McCord 2003).
It is important to understand that the value of an intervention for stakeholders might be different from the outcomes defined by the evaluator as the benchmark for “effectiveness”. The most important result of an alcohol prevention program might, for example, be its ability

(...) to cultivate understanding and support for alcohol policies, and (...) to motivate those who are at risk for hazardous or harmful alcohol use to seek help (WHO 2009, 54).

One explanation why so many interventions that have been shown to be ineffective by social researchers are still in use might simply be that they have other values to stakeholders than those that researchers are looking for.

One of the reasons for employers to work with alcohol prevention is because excessive alcohol consumption among employees “(...) can lead to accidents and injuries, a lowered production, increased health care costs, absenteeism and hang-over related performance problems (...)” (Eriksson Tinghög 2013, 12). While a program like PfL may not have a direct impact on the alcohol consumption among employees, it seems quite clear that the program could be used to create a shared “language” within a group. Increased risk awareness is one outcome that could be used by stakeholders as a way to generate good prospects for change in an alcohol culture. Shared perceptions of the risks associated with alcohol consumption, such as occupational risks, could increase awareness and the acceptance of alcohol policies and could potentially reduce the risk for organisational and individual harms caused by alcohol consumption (Sandberg 2011). The PfL evaluations indicate that it was not only risk awareness that increased among program participants, but also awareness of the organisation’s own alcohol policy. For example, prior to the implementation of the program within the Swedish Armed Forces 32 percent of those in the experiment group (N=446) and 36 percent of those in the control group (N=480) claimed that they were fairly or very familiar with the content of the organization’s alcohol policy. The difference between the groups was not statistically significant. Twelve months after the program, 54 percent of those in the experiment group claimed that they were fairly or very familiar with the content of the alcohol policy, compared to 42 percent in the control group. The difference between the groups was now statistically significant (p=<0.01).
Other outcomes that could be important for those involved in and affected by an activity could include changes that are not necessarily associated with alcohol consumption at all, such as improved cooperation between stakeholders, better functioning administration and care of students or employees, for example, improvements to people’s health or living conditions, enhanced documentation procedures or reduced bureaucracy within an organisation. However, a “one dimensional” experimental outcome evaluation with a focus on a limited set of outcome variables is not likely to pay much attention to such outcomes.

Both studies show that the possibility of using an alcohol prevention program to create a shared starting point for more systematic work with alcohol-related questions in an organisation is not necessarily dependent on whether or not the intervention has an impact on the participants’ alcohol behaviour. However, it is reasonable to assume that such a systematic prevention strategy effort requires that stakeholders agree about the usefulness of the intervention and that all involved parties participate in it (randomisation may no longer be an option!).

As much as I would like to claim that my evaluations of the PfL program display a desirable balance between the ambition to identify changes in alcohol behaviour among participants and recognizing other values, such as different program-related stakeholder perspectives and interests, the evaluations can be categorized foremost as quantitative outcome evaluations with a focus on a limited set of predefined desired outcomes. Nevertheless, simply having an open minded and critical approach to what the value and merits of the program were for different stakeholders has allowed me to better understand its value and merits on other levels than the predefined desired outcomes.

7.3 The top-down perspective

The “objective” and top-down perspective in outcome-oriented evaluations often fails to genuinely involve “ordinary people” (the non-professionals and non-experts) in the evaluation process.

Without dialogue, evaluation cannot serve democratizing interests, it cannot meaningfully assess if the public purposes of social investments have been realized (Greene 1999,168).
The process of talking to members of the Alcohol Committee, Prime for Life Instructors, students, university teachers and the departmental management at Örebro University opened up not only a deeper understanding of what was being evaluated and the context in which the PfL program was implemented, but also incorporated the stakeholders in the evaluation process. In a sense, the stakeholders were aware of the problems of implementation, and they were prepared for the results. Rather than waiting for the “verdict” on the effect of the program, both the Alcohol Committee and the Örebro University management were already actively involved in discussing different ways of developing their alcohol prevention strategies when the evaluation results were presented.

By contrast, the evaluation results from the Swedish Armed Forces study instead seem simply to have been shelved. Several years after the evaluation was presented there had still been no formal discussion within the Swedish Armed Forces managerial body about the results. Personally, I think that part of the reason for this is largely a of a pedagogic nature, and that I myself could have made a difference. Evaluation results that are both unexpected and negative (a category into which the evaluation results from the Swedish Armed Forces would fit) are often considered the ones that are most difficult for stakeholders to process (Torvatn & Rolfsen 2000). The likelihood that results will be “unexpected” is to a large extent dependent on the level of communication and interaction between the person who conducts the evaluation and those who requested it. In a top-down evaluation where quantitative data is collected with limited interaction between the evaluator and different stakeholders, it is often more difficult, unless certain measures are taken, to communicate important aspects of the evaluation, such as what criteria will be used to assess the intervention, methods used for data collection and what kind of results might be expected (Sandberg & Faugert 2012). The risk of adopting a “clinical” approach to an intervention, whereby interactions with different “subjects” are limited, is that those affected by the results may feel alienated from the evaluation process and that they are consequently more likely to reject the evaluation results – particularly if these results are negative. Negative and unexpected results are often actively resisted by stakeholders whereas results that are expected (regardless of whether they are positive or negative) are most often met with acceptance or satisfaction (Torvatn & Rolfsen 2000).
The evaluation of the PfL program within the Swedish Armed Forces was characterised, relatively speaking, by a greater distance between stakeholders and the evaluator than the Örebro University evaluation. Apart from the piloting of my evaluation questionnaires with a strategic sample of key stakeholders within the organisation, which was in part a measure intended to obtain acceptance for the evaluation, the evaluation design and data collection methods served to limit the level of interaction between myself as a researcher and those affected by the evaluation.

7.4 What could I have done differently?

The Örebro University study provided quite a broad, in-depth perspective on the implementation of the program; however, given the information about dissatisfaction with the program from departmental representatives, it would have been of value to further explore the underlying reasons for this. The interviews showed a complex relationship to alcohol at the department and a dichotomy in the way the “alcohol culture” was described, which I did not have a chance to fully investigate. Given this, it would have been interesting to go even deeper into understanding what this really was about. One hypothesis would be that departmental representatives reacted negatively towards the program because they were not involved in the process from the beginning and because they felt as if they had been chosen for the intervention because they had been identified by the Alcohol Committee and/or the Örebro University administration as being in special need of an alcohol prevention intervention. In order to understand their reaction to the PfL program, I think I could have more thoroughly investigated the factors that lead the Alcohol Committee and Örebro University to choose the Department of Hospitality, Culinary Arts and Meal Science as the venue for implementing the PfL program.

Further, it might have been possible to match the experiment group at the department with students from the same or a similar department, rather than with students in other programs on the main Örebro University campus. Although desirable from the perspective of producing a more solid statistical analysis (i.e. “better evidence”), I do not think that randomization of participants to the PfL program would have been possible without compromising the fact that the Alcohol Committee and Örebro University specifically
wanted to offer PfL as a voluntary program for all students at the Department of Hospitality, Culinary Arts and Meal Science.

The Swedish Armed Forces study was short on qualitative data that might have helped to better understand why the organisation was not capable of achieving its own goals in terms of both the program implementation and the data collection for the evaluation. The information available suggests that the attitude towards the Prime for Life program was not as positive at the regiments and units as it was at military headquarters. A more deliberate systematic study of the implementation process and of how the program was received at different levels within the organisation would have given me a better understanding of the context of Prime for Life in the Swedish Armed Forces.

Based on the experience of the evaluation of PfL in the Swedish Armed Forces, I think I should have designed a smaller outcome study set up as a randomized controlled experiment, rather than allowing the implementational goals of the Swedish Armed Forces to direct the evaluation. It would have been possible to randomly assign a smaller number of regiments with the capacity to conduct the Prime for Life program to either an experiment and control group and to personally take responsibility for collecting the pre- and post-test data, rather than having a such a large study population and being dependent on the regiments to administer the data collection process. A smaller study where I could personally have collected the data would have created natural arenas for communication with program managers and stakeholders during the evaluation, and it would also have taken some of the load off those regiments that felt overwhelmed by the project. Ideally, this would also have led to a study of higher statistical quality and perhaps the results would have been discussed and used more within the organisation.

My own interest in future studies of primary alcohol prevention would be to try to approach the question of what happens when high- and low-risk alcohol users are combined in the same group and participate in the same intervention. Does this create processes that are likely to support positive outcomes, or is it more likely that such a mix will obstruct the likelihood of achieving desired results? In a controlled experiment, it might be possible to screen participants’ alcohol use and then assign them to intervention groups with a) only high-risk consumers of alcohol, b) low-risk consumers of alcohol and c) groups with a mix of
high- and low-risk consumers. To gain a more in-depth understanding, I would preferably also conduct in-depth interviews about the program and its delivery with participants who have risky alcohol behaviour as well as with those who do not show any signs of risky alcohol use.

In the study of PfL in the Swedish Armed Forces, I think I was a bit naïve when it came to making sure that there were structured avenues for communication between myself as an evaluator and stakeholders before, during and after the evaluation. This may have had an impact on the extent to which the evaluation results were accepted, discussed and actually used by the Swedish Armed Forces. Since this study, I have become more aware of the necessity of planning evaluations so that the quality of an assessment includes strategies to improve communication with those who are affected by it before, during and after the evaluation.

My own role as a researcher and evaluator is not necessarily about striving to comply with the ideals of evidence-based practice; consequently not all of the reflections discussed here have dealt with ways to improve the quality of my studies from a statistical perspective. However, I think a higher level of control over program implementation and the possibility of randomizing individuals into experiment and control groups would have some benefits. It would not only have obtained “better evidence”, but would most likely also have applied techniques that would have simplified both data collection and analysis, at least in the Swedish Armed Forces evaluation.
8. Concluding remarks

My hope is that I have, in this thesis, been successful in both reporting my empirical studies and placing them in a context that has illustrated both their value and their shortcomings. The evidence-based practice movement constitutes an important backdrop to the thesis, along with the role played by experimental outcome evaluations in guiding decision-makers and practices in social policy in general, and in alcohol prevention in particular. While outcome studies can provide important information about “what works”, I have tried to illustrate the complex role of evaluations in policy-making and how a narrow outcome focus can diminish the value of understanding what is being evaluated and how it is supposed to work, the need for quality in implementation and the value of recognising different stakeholder perspectives. I will end by reconnecting to the role of evaluations in policy development as discussed in Chapter 2, and by expanding on my own thoughts on the possibilities for improving solutions to social problems by means of social science evaluations. Rather than adopting a linear approach to the relationship between science and policy-making, my own view fits into a more phenomenological understanding of social reality (Milani 2009).

I believe that evidence-based policy, just like other research policy fields, will develop and adopt new approaches to what constitutes “evidence”. This seems like a natural progression, since social policies are interpreted and delivered by complex and multi-faceted organisations. The rational process of decision making and the narrow focus on outcome variables that I have described in this thesis perhaps do not in themselves represent an entirely fair description of what evidence-based policy is, or what it will be in the future. The EBP movement may need to recognize that “evidence” is something that is socially mediated. A logical progression for EBP is perhaps be to shift from viewing innovation as a linear and technical process towards paying more attention to the development of new perspectives and tools that can improve opportunities for recognising different types of knowledge. The observations made by the National Board of Health and Welfare about the slow adoption of the ideas of evidence-based approaches into social policy and practice (see Chapter 2.4) have also been observed in other policy fields that have been influenced by the ideas of evidence-based medicine (see for example Walshe & Grundall 2001). In healthcare,
for example, tools that focus on concepts drawn from organisational and management sciences are today often used to improve the implementation of EBP (French et al. 2009). I would like to end this thesis by looking ahead to what might be needed to develop social interventions, regardless of whether or not they are based on the “evidence-based theorem”:

First, I think it is important to make a distinction between knowing what works and the ability to use what works: Alcohol policy in Sweden has been redirected from national policies to local prevention and increased individual responsibility. If the link between research and alcohol policy initiatives in our society was linear and research evidence actually guided practice, many alcohol prevention programs that are being used today would probably be “off the market”. As was discussed in Chapter 8, in addition to “primary outcomes”, organisations have many other values and interests that serve as an input in relation to prioritisations and decisions on whether to choose, abort or continue a certain alcohol prevention activity. “Political decisions about social policies are rarely the direct outcome of social science research” (Milani 2009, 44), and there are many other elements that go into the policy process (Weiss 1999). As was discussed in Chapter 2, the complexity of decision-making systems and the fact that policy-making is neither rational nor linear (Boaz & Hayden 2002) represents a major challenge to the ideas of EBP.

It is not sufficient to transfer evidence-based practices into a field in the absence of an understanding of what is needed to prepare decision-makers and practitioners to receive and implement this new knowledge (Barwick et al. 2005, 10). One way to increase the utilization of evidence locally might be for researchers to build stronger networks with a wide range of different stakeholders and to include them in the evaluation process. In addition to building a broader empirical knowledge-base about the merit and value of interventions, this might also create a sense of ownership of research findings among stakeholders. This could improve the transfer of knowledge to policy makers (Boaz & Hayden 2002). Such an approach places demands on researchers to develop skills that will enable

29 For example, being able to use “what works” in policy-making is dependent on how much an intervention costs. Few evaluators use cost-effectiveness measures as an indicator of how effective an intervention is (Brecher et al. 2005, Statskontoret 2002). At the same time, most governments around the world are looking for measures to reduce the cost of activities (House 2000).
them to see things from a stakeholder perspective and to understand the needs of policy
makers and practitioners.

Implementation has been relatively neglected in outcome research (Durlak 1998); few
published intervention-outcome evaluations refer to formal documentation that describes
the context, content and delivery of the intervention (Michie et al. 2009). The slow adoption
of evidence-based practice might in part be explained by the fact that there is a lack of
information on how to replicate evidence-based interventions (Dombrowski et al. 2007,
Michie et al. 2009, Riley et al. 2008). Interventions that work in one setting might not work in
program is well documented, manual-based and executed by well-trained instructors (see
Chapter 7.1.3). However, I think that both of my evaluations clearly illustrate that the
implementation of the program (in terms of gaining approval for the program, why it was
chosen and how it was introduced to participants) was not without its problems. These
issues regarding implementation may not have affected the primary program outcomes per
se, but they may have had other negative consequences. This might for example be serving
to limit the possibilities for decision-makers to make a long-term commitment to the
intervention and similar activities and for stakeholders to use the alcohol prevention
program as a tool to develop a structured basis for alcohol policy and prevention.

In order to be able to move the development of social interventions to a higher level, at
which effective interventions can be understood and implemented successfully, I believe
that there is a need to broaden the perspective both on what constitutes “evidence” and on
how evaluations should be performed. A broader and more systematic approach to
evaluation should in my opinion include:
1. Thorough descriptions of both the intervention and the context in which it is being implemented so that the intervention can be assessed in a fair way and so that knowledge can be disseminated about which conditions and elements are supportive of effective interventions.

2. Evaluations that include studies of the quality of implementation, in order to improve our understanding of the prerequisites necessary for implementing effective interventions.

3. The use of a wide range of outcome measures that acknowledge the different values and interests associated with social interventions, not only “primary measures” pre-defined by the evaluator or by vague political goals.

4. Active contacts with stakeholders to capture a wide range of perspectives on the intervention and its merits and values. This is important in order to obtain acceptance for evaluations and to make sure that evaluation results are accepted and used.

5. A genuine utilization of method triangulation in order to increase validity and so that a broad set of values of an activity can be addressed and assessed from different perspectives.

It is important to remember that public management in Sweden is to a large extent result-oriented and that this has increased the focus on measuring results and indicators of quality within the public sector (Lindgren 2008). An increased supply of outcome studies focusing on the effectiveness of interventions is crucial to the implementation of evidence-based practice. Having said this, there is – generally speaking – a political demand for a narrower, rather than a broad, approach to evaluations of social interventions.
Perhaps the greatest challenge for evaluators of alcohol prevention is to identify the contextual circumstances that make good interventions effective, and to find out how these mechanisms can be transferred to other situations. An approach where the aim is to find out how and under what conditions a given measure will produce its impact is sometimes referred to as realistic evaluation (Tilley 2000). Such an approach would represent an attempt to apply a broad view of evaluation that goes beyond a narrow outcome perspective and that includes stakeholder expectations and a focus on the quality of implementation. Rather than relying on experimental designs and quantitative data, the use of multiple evaluation designs, and multiple data sources and data collection methods (i.e. data and methodological triangulation) can be used to increase both the breadth and depth of our understanding of social interventions, and also to increase study accuracy (Jick 1979, Hussein 2009). This, I believe, could also improve the possibilities of reproducing and implementing “what works” in an effective way. In the long run, the EBP movement can most probably only be successful if it acknowledges the need to understand how and in what circumstances things (might) work.
References


Ramstedt, M. et al. (2010). *Tal om alkohol 2010. En statistisk årsrapport från monitor-
projektet. Research Report #60. Centre for Social Research on Alcohol and Drugs, Stockholm University, Stockholm.


WHO (2004): *What are the most effective and cost-effective interventions in alcohol control?* World Health Organization Regional Office for Europe, Copenhagen, Denmark.


# Appendix 1: The AUDIT-test

Babor et al (2010)

<table>
<thead>
<tr>
<th>Questions</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How often do you have a drink containing alcohol?</td>
<td>Never</td>
<td>Monthly or less</td>
<td>2-4 times a month</td>
<td>2-3 times a week</td>
<td>4 or more times a week</td>
</tr>
<tr>
<td>2. How many drinks containing alcohol do you have on a typical day when you are drinking?</td>
<td>1 or 2</td>
<td>3 or 4</td>
<td>5 or 6</td>
<td>7 to 9</td>
<td>10 or more</td>
</tr>
<tr>
<td>3. How often do you have six or more drinks on one occasion?</td>
<td>Never</td>
<td>Less than monthly</td>
<td>Monthly</td>
<td>Weekly</td>
<td>Daily or almost daily</td>
</tr>
<tr>
<td>4. How often during the last year have you found that you were not able to stop drinking once you had started?</td>
<td>Never</td>
<td>Less than monthly</td>
<td>Monthly</td>
<td>Weekly</td>
<td>Daily or almost daily</td>
</tr>
<tr>
<td>5. How often during the last year have you failed to do what was normally expected of you because of drinking?</td>
<td>Never</td>
<td>Less than monthly</td>
<td>Monthly</td>
<td>Weekly</td>
<td>Daily or almost daily</td>
</tr>
<tr>
<td>6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?</td>
<td>Never</td>
<td>Less than monthly</td>
<td>Monthly</td>
<td>Weekly</td>
<td>Daily or almost daily</td>
</tr>
<tr>
<td>7. How often during the last year have you had a feeling of guilt or remorse after drinking?</td>
<td>Never</td>
<td>Less than monthly</td>
<td>Monthly</td>
<td>Weekly</td>
<td>Daily or almost daily</td>
</tr>
<tr>
<td>8. How often during the last year have you been unable to remember what happened the night before because of your drinking?</td>
<td>Never</td>
<td>Less than monthly</td>
<td>Monthly</td>
<td>Weekly</td>
<td>Daily or almost daily</td>
</tr>
<tr>
<td>9. Have you or someone else been injured because of your drinking?</td>
<td>No</td>
<td>Yes, but not in the last year</td>
<td>Yes, during the last year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Has a relative, friend, doctor, or other health care worker been concerned about your drinking or suggested you cut down?</td>
<td>No</td>
<td>Yes, but not in the last year</td>
<td>Yes, during the last year</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2: Questionnaire (T2) Swedish Armed Forces

Enkät (12-månadersuppföljning, T2)

Centrum för Socialvetenskaplig forskning om alkohol- och droger (SoRAD) vid Stockholms universitet har fått av Försvarsmakten i uppdrag att genomföra en studie i syfte att utvärdera det drogpreventiva arbetet inom organisationen. För ca ett år sedan genomgick du en utbildning i Prime for Life och fick i anslutning till denna besvara en enkät angående alkoholvanor. Denna enkät följer upp den studien.


Ec1. Vänligen ange vilket år och månad Du genomför denna enkät

Ex: 0 6 0 1 (januari 2006)

Ec2. Dra bort siffran för Din födelsemånad från de fyra sista siffrorna i Ditt personnummer. Skriv summan i rutorna nedan. Exempel: 741113-7110 \( \rightarrow \) 7110 - 11 =7099

☐ ☐ ☐ ☐

Din identitet kan inte avslöjas utifrån ”koden” ovan. Uppgiften används för att vi skall få en något så när unik ”kod” för varje deltagare och därmed kunna kontrollera att samma personer deltagit i delstudierna.
Bakgrundsfrågor

Ec3. Är Du...

- [ ] Kvinna
- [ ] Man

Ec4. När är Du född?

- [ ] Före 1940
- [ ] 1946-1950
- [ ] 1956-1960
- [ ] 1966-1970
- [ ] 1976-1980
- [ ] 1940-1945
- [ ] 1951-1955
- [ ] 1961-1965
- [ ] 1971-1975
- [ ] Efter 1980

Ec5. Vilken typ av tjänst har Du inom Försvarsmakten?

- [ ] Officerare
- [ ] Civilanställd

Ec6. Vilket vapenslag tillhör Du?

- [ ] Flygvapnet
- [ ] Armén
- [ ] Marinen
- [ ] Annat (t.ex. FM-log, HKV)
Frågor om alkohol

Ec7. Dricker Du alkohol?

☐ Nej → gå till fråga Ec24

☐ Ja

Ec8. Hur ofta har Du under de senaste 2 veckorna druckit någon form av alkohol?

☐ Aldrig → gå till fråga Ec11

☐ En gång

☐ 2-3 gånger

☐ 4-7 gånger

☐ I stort sett dagligen (8-12 gånger)

☐ Dagligen

Ett "glas alkohol" motsvarar

1 flaska (33cl) starköl/cider eller

1 burk (50 cl) folköl eller

Ec9. Vilket är det högsta antalet glas (se ovan) Du druckit vid ett tillfälle under de senaste 2 veckorna?

☐ 1-2 ☐ 5-6 ☐ 10-12 ☐ 16-18 ☐ 22 eller fler

☐ 3-4 ☐ 7-9 ☐ 13-15 ☐ 19-21
Ec10. Vid hur många tillfällen under de senaste 2 veckorna har Du druckit ungefär så många glas som Du angav i frågan ovan (Ec9)?

☐ 1-2 ggr  ☐ 5-6 ggr  ☐ 9-10 ggr  ☐ 13-14 ggr
☐ 3-4 ggr  ☐ 7-8 ggr  ☐ 11-12 ggr

Ec11. Hur ofta har Du under de senaste 12 månaderna druckit någon form av alkohol?

☐ Aldrig → gå till fråga Ec22

☐ En gång i månaden eller mindre

☐ 2-4 gånger i månaden

☐ 2-3 gånger i veckan

☐ 4 ggr i veckan eller mer

Ec12. Hur många glas (se ovan) konsumerar Du under en typisk dag då Du dricker alkohol?

☐ 1-2  ☐ 5-6  ☐ 10-12  ☐ 16-18  ☐ 22 eller fler
☐ 3-4  ☐ 7-9  ☐ 13-15  ☐ 19-21
Ec13. Hur ofta dricker Du sex glas (se ovan) alkohol eller mer vid samma tillfälle?

☐ Aldrig

☐ Mer sällan en gång i månaden

☐ Varje månad

☐ Varje vecka

☐ Dagligen eller nästan varje dag

Ec14. Hur ofta under det senaste 12 månaderna har Du funnit att Du inte kunnat sluta att dricka alkohol när Du väl börjat?

☐ Aldrig

☐ Mer sällan en gång i månaden

☐ Varje månad

☐ Varje vecka

☐ Dagligen eller nästan varje dag
Ec15. Hur ofta under de senaste 12 månaderna har Du låtit bli att göra det Du skulle ha gjort därför att Du drack alkohol?

☐ Aldrig

☐ Mer sällan en gång i månaden

☐ Varje månad

☐ Varje vecka

☐ Dagligen eller nästan varje dag

Ec16. Hur ofta under de senaste 12 månaderna har Du behövt inta alkohol på morgonen för att komma igång efter att ha druckit mycket alkohol dagen före?

☐ Aldrig

☐ Mer sällan en gång i månaden

☐ Varje månad

☐ Varje vecka

☐ Dagligen eller nästan varje dag
Ec17. Hur ofta under de senaste 12 månaderna har Du haft skuldkänslor eller varit ångerfull efter att ha druckit alkohol?

- Aldrig
- Mer sällan en gång i månaden
- Varje månad
- Varje vecka
- Dagligen eller nästan varje dag

Ec18. Hur ofta under de senaste 12 månaderna har det varit omöjligt för Dig att komma ihåg vad som hände kvällen innan därför att Du druckit alkohol?

- Aldrig
- Mer sällan en gång i månaden
- Varje månad
- Varje vecka
- Dagligen eller nästan varje dag
Ec19. Har Du eller någon annan blivit skadad som ett resultat av Ditt drickande?

☐ Nej

☐ Ja, men inte under det senaste året

☐ Ja, under det senaste året

Ec20. Har närstående, vänner, läkare eller någon annan inom sjukvården varit oroad över Dina alkoholvanor eller föreslagit att Du bör minska Din alkoholkonsumtion?

☐ Nej

☐ Ja, men inte under det senaste året

☐ Ja, under det senaste året

Ec21. Om Du tänker på de senaste 12 månaderna, har Du vid något tillfälle känt att Din alkoholkonsumtion har en positiv inverkan på...

<table>
<thead>
<tr>
<th></th>
<th>Nej</th>
<th>Ja</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) dina vänkapsrelationer eller ditt sociala liv?</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>b) din fysiska hälsa?</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>c) ditt familjeliv eller ditt äktenskap/samboförhållande</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>d) din ekonomi?</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>e) ditt arbete?</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
Ec22. Skulle Du vilja ändra Din alkoholkonsumtion?

☐ Nej, jag dricker så mycket eller lite som jag är nöjd med

☐ Ja

☐ Vet ej

Ec23. Hur stor risk bedömer Du Dig själv ha att utveckla alkoholproblem?

☐ Ingen risk alls

☐ Ganska liten risk

☐ Ganska stor risk

☐ Mycket stor risk

☐ Jag har redan alkoholproblem

☐ Vet ej

Ec24. Har Du under de senaste 12 månaderna deltagit i någon utbildning som syftat till att öka kunskapen om, eller påverka, Din alkoholkonsumtion?

☐ Nej

☐ Ja, Prime for Life

☐ Ja, annan utbildning

☐ Vet ej
Ec25. Hur väl insatt är Du i innehållet i Försvarsmaktens drogpolicy ("policy beträffande missbruk av alkohol, narkotika och dopningsmedel")?

- Visste ej att det fanns en policy
- Inte alls
- Något
- Ganska väl
- Mycket väl

Ec26. Hur väl stämmer dessa påståenden med vad Du innerst inne tror?

<table>
<thead>
<tr>
<th>Stämmer mycket</th>
<th>Stämmer ganska</th>
<th>Stämmer ganska</th>
<th>Stämmer mycket</th>
<th>Vet ej</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ganska</td>
<td>Dåligt</td>
<td>Bra</td>
<td>Bra</td>
<td>Dåligt</td>
</tr>
</tbody>
</table>

a) Vem som helst kan utveckla alkoholism

b) Hur ofta man dricker påverkar risken för att utveckla alkoholism

c) Hur mycket man dricker påverkar risken för att utveckla alkoholism

d) Den mängd alkohol jag själv konsumerar har en negativ inverkan på mitt liv
e) Att tåla mycket alkohol är ett tecken på ökad risk för att utveckla alkoholism

f) Vi föds alla med en alkohol-tolerans som bestäms av våra arvsanlag

g) Varje gång man dricker tills man blir påverkad och sedan lite till så ökar toleransen

h) Att dricka mycket alkohol vid ett tillfälle innebär större risk för problem än att dricka mindre mängder ofta

i) Om andra drack samma mängd alkohol som jag skulle det ha negativ inverkan på deras liv

Frågor om din livssituation

Ec27. Ange vilka av nedanstående påståenden som Du tycker stämmer in på vad som hänt i Ditt liv under de senaste 12 månaderna?

<table>
<thead>
<tr>
<th></th>
<th>Nej</th>
<th>Ja</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Fått barn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Blivit sambo/gift</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
c) Skiljd/separerat

 d) Uppsagd från jobbet

 e) Drabbats av våldsbrottbrott

 f) Mått psykiskt dåligt

 g) Varit långtidssjukskriven

 h) Dömts för brott

 i) Använt narkotika

 j) Vantrivts på arbetet

 k) Flyttat från tätort/stad ut på landet

 l) Flyttat från landet till tätort/stad

 m) Fått försämrad ekonomi

 n) Fått förbättrad ekonomi

 o) Bytt umgångeskrets

 p) Fått stöd eller behandling för alkohol-problem

 q) Känt oro inför förändrad arbetssituation

 Tack för Din medverkan!
## Appendix 3: Outcomes of Prime for Life, Swedish Armed Forces

### TABLE 3.1: EFFECTS OF PRIME FOR LIFE (PFL) ON KEY OUTCOMES 12 MONTHS AFTER THE PROGRAM (T2). RESULTS FROM ANALYSES OF COVARIANCE (ANCOVA), ADJUSTING FOR PRE-TEST DIFFERENCES IN OUTCOME, AGE AND MILITARY BRANCH (FULL SAMPLE, 95% CONFIDENCE INTERVALS IN PARENTHESES). MEN, SWEDISH ARMED FORCES.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Risk awareness</th>
<th>No. of drinks on a typical day drinking</th>
<th>How often alcohol consumption</th>
<th>Frequency binge drinking</th>
<th>Total AUDIT score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
</tr>
<tr>
<td>PFL</td>
<td>5.58</td>
<td>[4.82 – 6.33]</td>
<td>1.16</td>
<td>[1.01 – 1.31]</td>
<td>-0.12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of observations</th>
<th>( n_{\text{total}} )</th>
<th>( n_{\text{PFL}} )</th>
<th>( n_{\text{control}} )</th>
<th>( n_{\text{total}} )</th>
<th>( n_{\text{PFL}} )</th>
<th>( n_{\text{control}} )</th>
<th>( n_{\text{total}} )</th>
<th>( n_{\text{PFL}} )</th>
<th>( n_{\text{control}} )</th>
<th>( n_{\text{total}} )</th>
<th>( n_{\text{PFL}} )</th>
<th>( n_{\text{control}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>729</td>
<td>341</td>
<td>388</td>
<td>712</td>
<td>335</td>
<td>377</td>
<td>714</td>
<td>336</td>
<td>378</td>
<td>709</td>
<td>331</td>
<td>378</td>
</tr>
</tbody>
</table>

SMD=Standardized mean difference (Cohen’s d).
Table 3.2: Effects of Prime for Life (PFL) on key outcomes 12 months after the program (T2). Results from analyses of covariance (ANCOVA), adjusting for pre-test differences in outcome, age and military branch (full sample, 95% confidence intervals in parentheses). Women, Swedish Armed Forces.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Risk awareness</th>
<th>No. of drinks on a typical day drinking</th>
<th>How often alcohol consumption</th>
<th>Frequency binge drinking</th>
<th>Total AUDIT score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
<td>SMD</td>
<td>Beta-coef.</td>
</tr>
<tr>
<td>PFL</td>
<td>4.39</td>
<td>0.91</td>
<td>-0.04</td>
<td>-0.07</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>[4.87 – 5.91]</td>
<td>[0.62 – 1.20]</td>
<td>[-0.18 – 0.10]</td>
<td>[-0.38 – 0.24]</td>
<td>[-0.13 – 0.14]</td>
</tr>
</tbody>
</table>

| No. of observations | n<sub>total</sub>= 182 | n<sub>PFL</sub>= 98 | n<sub>control</sub>= 84 | n<sub>total</sub>= 171 | n<sub>PFL</sub>= 91 | n<sub>control</sub>= 80 | n<sub>total</sub>= 177 | n<sub>PFL</sub>= 93 | n<sub>control</sub>= 84 | n<sub>total</sub>= 171 | n<sub>PFL</sub>= 87 | n<sub>control</sub>= 75 |

SMD=Standardized mean difference (Cohen’s d).
Kriminologiska institutionens rapportserie

1997:1 Ringman, Karin, Brottsförebyggande åtgärder i en förort.
1998:1 Backström, Malin (red.), Homosexualitet i dag.
2001:1 Westfelt, Lisa, Organisatoriska förändringar inom svensk polis under 1900-talet med avseende på kampen mot organiserad brottslighet.
2001:3 Shannon, David, Graffiti and Adolescent Delinquency. An Analysis of Short Term Career Trajectories.
2003:1 Tham, Henrik (red.), Forskare om narkotikapolitiken.
2005:1 Skrinjar, Monica, Bilder av och åtgärder mot narkotika(miss)brukare – en reflexiv studie.
2005:2 Pettersson, Tove, Polisingripanden vid eget bruk av narkotika. Särbehandlade personer med utländsk bakgrund?
2005:3 Heber, Anita, Var rädd om dig! En litteraturöversikt om rädslan för brott.
2006:1 Litzén, Staffan, Oro för brott i urban miljö. Trygghetsundersökningar med anknytning till Stockholm.
2009:1 von Hofer, Hanns (red.), Leif Lenke in memoriam
2010:1 Jerre, Kristina & Tham, Henrik, Svenskarnas syn på straff.
2011:1 Kardell, Johan, Utländsk bakgrund och registrerad brottslighet – överrepresentationen i den svenska kriminalstatistiken.
2012:1 Jonsson, Adam, ”Man hör aldrig om killar som blir sexuellt trakasserade”. En rapport om sexuella trakasserier i skolan.
2013:1 Sandberg, Bo, *Looking beyond what doesn’t work – Results, policy implications and lessons learned from two outcome evaluations of the Prime for Life alcohol prevention program.*